FOOT DROP AND FATIGABILITY IN PEOPLE WITH MULTIPLE SCLEROSIS

GEORGIA ANDREOPOULOU

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

QUEEN MARGARET UNIVERSITY

2019

‘Nexus – the search for’ 2018

This work was inspired by a research study into physical performance related fatigue in physically active MS patients, before, during and after exercise. Also the evolution of motor fatigability over time. Depicted in the searing landscape is a figure representing the tireless search for a connection – a clue – and answer to the many questions agonized by researchers, who are hardly ever seen. Behind the image are a trail of computerised images, similar to the images, utilised to assist accurate patient measurement of movement.

B.J. Leslie, MS participant

Abstract

People with Multiple Sclerosis (pwMS) often experience walking impairments such as foot drop which can lead to trip and falls. Foot drop can be either transient and is often induced by exercise (fatigability) in pwMS whose walking ability is not affected and can become more fixed with disease progression. The overall aim of this PhD was to explore foot drop, its presence in pwMS with different disability levels and the psychometric properties of outcomes used to evaluate walking impairments. The first study in this thesis was a systematic review into the level of evidence for the psychometric properties of walking measures that have been used to evaluate the effect of assistive technology such as FES for foot drop in MS. Moderate to strong psychometric evidence was found for the Multiple Sclerosis Walking Scale, Timed 25 Foot Walk, 6 minute and 10 meter walk tests. There were no psychometric studies for three-dimensional (3D) gait kinematics in pwMS even though it was one of the most frequently used outcome measures. The second study assessed the test-retest reliability of 3D ankle kinematics and spatiotemporal parameters in pwMS, with low Expanded Disability Status Scale (EDSS < 3.5) and in those with moderate to high EDSS (EDSS: 4-6). Reliability was excellent for ankle kinematics and spatiotemporal parameters in both groups, with lower minimal detectable change ($MDC_{95\%}$) values in the low EDSS group compared to the higher EDSS group. The third study investigated transient exercise induced foot drop in highly physically active pwMS (EDSS < 3.5) and health controls using 3D kinematics. It was found that 6 out of 15 pwMS and none of the healthy controls presented this phenomenon. The fourth study examined the direct orthotic effect of FES during dual-tasking (i.e. walking combined with a cognitive task) and after inducing fatigability. Low to moderate effect sizes indicated that the direct orthotic effect was higher under dual-task and fatiguing conditions but this needs to be confirmed in appropriately powered studies.

In conclusion, the studies in this thesis have contributed to the psychometric evidence of gait kinematics in pwMS, have objectively documented the presence of transient foot drop in highly physically active pwMS and orthotic effect of FES under a variety of conditions simulating the perceived benefits in 'real life' reported by FES users.

Key words: multiple sclerosis, foot drop, fatigability, FES, psychometric properties, 3D gait kinematics

Acknowledgements

This has been a long journey with many memorable good times, but also a few difficult ones, that would have not existed without the people that have been supporting me throughout. There are so many people that I would like to thank even though I cannot name everybody here.
Firstly, I would like to thank my supervisory team dr. Marietta van der Linden and prof. Tom Mercer. I learnt so much from them, and their constant support and motivation helped me to arrive at this stage. Especially Marietta, who is the person I admire the most, has always been there for my happy moments, understood, and encouraged me in my dark moments of this journey. Her constant positivity and knowledge made me keep going, always been there even late at nights! I could keep going on for her in many pages!
I am grateful for the bursary that I have received for three years from QMU and the funding from the Anne Rowling Clinic in the final year of this PhD. Especially Dr. Don Mahad, who has believed in me and supported me throughout all these years. I have learnt so much from him with our talks and I am so grateful of having the chance to work with him.
These four years I had the chance to meet some amazing people that have been part of my life. Suzanne and Christos have been there for me from the start until the end, literally the end! Waited with me until 2am to make sure I was not alone as I was waiting to be prepared for the theatre and my surgery. We had of course so many celebration times and will have many more in the future! I would like to thank Pelly for hearing me moaning and encouraging me to understand myself better with our long conversations and for just been there for me. I would also like to thank Ashley, which is one of the best persons I have met here in Scotland and how lucky I am to have had the chance to live with her, always there for me!
Well, at this point I could not leave out my friends in Greece that always supported my crazy ideas and ventures and no matter how far I was always been there for me. Especially, Christiana, Konstantina, Dimitra, Giannis and Alekos for their constant support! They have been there for me since I remember myself and if I will start now with our stories I will need another thesis to write!
I would not be at this stage now without the support of my parents, Paschalis and Vasiliki, that all these years they have done anything but supporting me and believing in me. In these four years, I became an auntie and wee Antonis is the person that I have been speaking with every day and gave me so much strength to continue with this journey. My sister Eleni has always been the rock, always believing and encouraging me to achieve my dreams.
Lastly but not least, this PhD could not have been completed without the precious time that the participants put into it. I am so grateful for meeting each one of them, our chats during the hours that we spent together and how much I learnt from them! The positive attitude of all these people that even a chronic disease cannot stop you from doing what you love, changed my whole attitude towards life.

Contents

List of tables

List of figures

Abbreviations

10mWT	10-meter walk test
2MWT	2-minute walk test
3DGA	Three-dimensional gait analysis
6MWT	6-minute walk test
9-HPT	Nine-Hole Peg Test
AFOs	Ankle Foot Orthoses
ARC	Anne Rowling Clinic
ASIS	Anterior superior iliac spine
B&A	Bland-Altman
BMI	Body Mass Index
CFQ	Chalder Fatigue Scale
CIS	Clinically isolated syndrome
CIS-20R	Checklist Individual Strength questionnaire
CMC	Coefficient of Multiple Correlation
CMI	Cognitive motor interference
CNS	Central nervous system
CONSORT	Consolidated Standards of Reporting Trials
COSMIN	COnsensus-based Standards for the selection of health Measurement INstruments
CP	Cerebral palsy
DF	Dorsiflexion
D-FIS	Daily Fatigue Impact Scale
DGI	Dynamic Gait Index
DMT	Disease modifying therapy
DT	Dual-task
DTC	Dual-task cost
DWI	Distance walk index
EDSS	Expanded Disability Status Scale
EMG	Electromyography
EMIF-SEP	French adaption of Fatigue Impact Scale

ES	Effect size
FAI	Fatigue Assessment Inventory
FAI	Frenchay Activities Index
FDS	Fatigue damage spectrum
FES	Functional Electrical Stimulation
FIS	Fatigue Impact Scale
FKS	Fatigue index Kliniken Schmieder
FS	Feeling Scale
FS	Functional system
FSMC	Fatigue Scale for Motor and Cognitive Function
FSS	Fatigue Severity Scale
GDI	Gait Deviation Index
GMFCS	Gross Motor Function Classification System
GPS	Gait Profile Score
GRI	Guyatt Responsiveness Index
GVS	Gait Variable Scores
HADS	Hospital Anxiety and Depression Scale
HAI	Hauser Ambulation Index
HR-PROs	Health related patient-reported outcomes
IC	Initial contact
ICC	Intra-Class correlation coefficient
ICF	International Classification of Functioning
IQR	Interquartile range
IVMP	Intravenous methylprednisolone
KAD	Knee alignment device
LA	Least affected
LoA	Limits of agreement
MA	Most affected
MCID	Minimal clinically important difference
MDC	Minimal detectable change
MFI	Multidimensional fatigue inventory
MFIS	Modified Fatigue Impact Scale

MFSS	Multiple sclerosis-specific fatigue severity scale
MIC	Minimal Important Change
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
MSFC	Multiple Sclerosis Functional Composite
MSIS-29	Multiple Sclerosis Impact Scale-29
MSNQ	Multiple Sclerosis Neuropsychological Screening Questionnaire
MSWS-12	12-Item MS Walking Scale
NFI-MS	Neurological Fatigue Index
NHS	National Health Services
PAR-Q+	Physical Activity Readiness Questionnaire for Everyone
PASAT	Paced auditory serial addition test
PCI	Physiological Cost Index
PDDS	Patient Determined Disease
PI	Principal investigator
PIS	Participant Information Sheet
PP	Primary progressive
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PS-F	Performance Scale Fatigue
PSIS	Posterior superior iliac spine
pwMS	people with MS
QoL	Quality of life
RMS	Root mean square
ROC	Receiver operating characteristics
RPE	Rate of perceived exertion
RR	Relapsing-remitting
RT/MT	Reaction time/movement time
RVGA	Rivermead Visual Gait Assessment
SD	Standard deviation
SDC	Smallest detectable change
SEM	Standard error of measurement

SMD	Standardised Mean Differences
SOFI	Swedish occupational fatigue inventory
SP	Secondary progressive
SRM	Standardised response mean
SRT	Shuttle run test
SSWS	Self-selected walking speed
T100mW	Timed 100 meter walk
T25FW	Timed 25 foot walk
T500mW	Timed 500 meter walk
TUG	Timed up and go test
U-FIS	Unidimensional Fatigue Impact Scale
VAS	Visual Analogue Scale
VO_2 peak	Peak oxygen uptake
VR	Virtual-reality
WEIMUS	Würzburg Fatigue Inventory for Multiple Sclerosis
WLG	Word list generation

Publications/presentations arising from this thesis

Andreopoulou G., van der Linden M.L., Mercer T.H., 2017. 22nd Annual RIMS Conference. Psychometric properties of walking performance and lower limb function measures used to assess effects of assistive technology to treat foot drop in multiple sclerosis – A COSMIN review Poster presentation, Barcelona, 4-6 May 2017.

Andreopoulou G., Mercer T.H., van der Linden M.L., 2018. Walking measures to evaluate assistive technology for foot drop in multiple sclerosis: A systematic review of psychometric properties. *Gait & Posture*, 61: 55-66.

Andreopoulou G., Mahad D.J., van der Linden M.L., Mercer T.H., 2018. 21st SIG Mobility RIMS Meeting. Deterioration of gait kinematics in highly active people with MS after an exercise running task – A preliminary analysis. Oral presentation, Glasgow, 28-29 September 2018.

Andreopoulou G., Mercer T.H., Mahad D.J., van der Linden M.L., 2019. MS Frontiers. Deterioration of gait kinematics in minimally impaired people with MS after an exercise-running/walking task. Oral presentation, Bath, 4-5 July 2019. Funded by Rosemary Ann Price Award

Andreopoulou G., Mahad, D.J., Mercer T.H., van der Linden M.L., 2019. Test-retest reliability and minimal detectable change of ankle kinematics and spatiotemporal parameters in MS population. *Gait & Posture,* Accepted manuscript.

Related publications/presentations

van der Linden M.L., Andreopoulou G., Campbell B., Hooper J.E., Scopes J., Mercer T.H., 2017. 14th Congress of the European Forum for Research in Rehabilitation. Changes in ankle kinematics over the duration of a six minute walk test in people with multiple sclerosis. Poster presentation, Glasgow, 24-27 May 2017.

Van der Linden M.L., Andreopoulou G., Scopes J., Hooper J.E., Mercer T.H., 2018. Ankle kinematics and temporal gait characteristics over the duration of a 6-minute walk test in people with multiple sclerosis who experience foot drop. *Rehabilitation Research and Practice*, 2:2018

Chapter 1. Introduction

Multiple sclerosis (MS) is chronic inflammatory demyelinating disease that affects the central nervous system (CNS) and typically strikes adults (McDonald & Sears, 1970). Usually it presents between the ages of 20-40 and it affects 2.3 million people worldwide with twice as many women affected compared to men (Browne et al, 2014). According to the Scottish MS Register, 425 people diagnosed with MS in 2017 and it is estimated that the incidence for Scotland is 8.6 per 100,000/year (Scottish MS Register, viewed online 2018). Multiple sclerosis affects people in different ways, with people experiencing a variety of symptoms. The most commonly reported symptoms are fatigue, weakness and movement disturbances (Crayton & Rossman, 2006). In a study by Hessen et al. (2008) gait function was most frequently rated as the most important domain by pwMS.

The Expanded Disability Status Scale (EDSS) is a clinician-administered scale that assesses the neurological disability in MS population and describes the disease progression based on an ordinal system ranging from 0 (normal neurological status) to 10 (death due to MS) (Kurtzke, 1983). Although increased severity of walking impairments is associated with a longer disease duration and higher EDSS, small gait deficits have also been observed in people who are relatively mildly affected by MS (Benedetti et al, 1999; Nogueira et al, 2013).

One common walking impairment is foot drop (drop-foot or dropped foot), which is the decreased ankle dorsiflexion during the swing phase of gait and can result in trips and falls. The prevalence of foot drop has not been reported in peer-

reviewd literature, but based on the records of the MS Therapy Centre Lothian it is estimated that 48% of the pwMS that visit the centre experience foot drop. Further, there are anecdotal reports of foot drop even at early stages of the disease. This type of foot drop often comes on during intense or long duration physical activities such as walking or running, but recovers after a period of rest. To the author's knowledge, this type of foot drop has not been objectively documented for example by gait analysis.

Fixed foot drop is commonly experienced by people with EDSS of 4 onwards and is often treated by Ankle Foot Orthoses (AFOs) or Functional Electrical Stimulation (FES). The use of FES has gained popularity over the latest decade, with some pwMS reporting benefits including improvement of gait, reduced fatigue and reduced tripping and falling (Taylor et al, 1999; Bulley et al, 2015; (Miller) Renfrew et al, 2018). Another important benefit of FES reported by pwMS is that it reduces the mental effort of walking and as a result less concentration is needed to walk (Bulley et al, 2015; (Miller) Renfrew et al, 2018). This is of great importance, since most of the daily activities include performing a motor and a cognitive task at the same time. The execution of a motor task simultaneous with a cognitive task is termed dual-tasking and studies have indicated that when performing a cognitive task together with a walking task, pwMS exhibited deterioration in gait performance. However, this perceived benefit of FES use while performing a cognitive task has not been documented objectively.

There is wide variety of outcome measures to assess walking performance, for example with and without FES, in clinical and research settings for the MS

population, such as self-reported questionnaires, clinician-assessed rating scales and short or long distance walking performance tests and three-dimensional (3D) gait analysis. The latter is often considered the 'gold' standard as it can detect even subtle changes. In order to allow for meaningful interpretation of the results, the outcome measures should be reliable, valid and responsive to change, as well as practical and appropriate for each setting. Consequently, there is an increase on the number of studies investigating the level of evidence for the psychometric properties of outcome measures and specific guidelines have been developed [i.e. COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN)] in order to evaluate and summarize the level of evidence of outcome measures.

This thesis will focus on foot drop and treatment of foot drop using FES in pwMS. The following Chapter 2 will therefore first describe the definition, prevalence and symptoms of MS. This is then followed by a detailed report on the walking impairments and especially foot drop (both transient and fixed) in pwMS with different walking abilities and treatment options for foot drop such as FES. In addition, this literature review includes a description of the psychometric properties of outcome measures of walking performance in pwMS, along with the current knowledge on dual-task performance in the MS population. Both topics are of importance for the assessment of the impact of FES to treat foot drop in pwMS. Finally, at the end of Chapter 2 the overall aim, objectives and research questions of this thesis will be presented.

Chapter 2. Literature review

2.1 Purpose of the chapter

This chapter will present an overview of multiple sclerosis (MS) in terms of its clinical definition, pathophysiology, incidence and prevalence and symptoms focusing mostly in performance related fatigue versus perceived fatigue. Further, in this chapter gait impairments which are common in the MS population, such as foot drop, will be explored as well as assistive technology to treat foot drop [Functional electrical stimulation (FES)]. Another part of the literature review will describe those outcome measures that have been used to assess walking performance in the MS population and the importance and need for psychometrically robust outcomes in this population will be highlighted. The next part will provide a background on dual tasking (i.e. performing motor and cognitive tasks simultaneously) in people with MS (pwMS) and how it affects walking performance. At the end of this literature review, gaps in evidence identified will be highlighted and the rationale of the four studies that form this thesis will be explained.

2.2 Multiple Sclerosis – An overview

2.2.1 Multiple sclerosis and pathology

Multiple sclerosis (MS) was defined more than a century ago and is an immune-mediated demyelinating disease of the central nervous system (CNS). It is manifested with the presence of focal areas of inflammatory-mediated demyelination of the brain and spinal cord (Trapp et al, 2008). Early in the 19th century Robert Carswell (1838) and Jean Cruveilhier (1841) macroscopically

recorded scarring or lesions on the brain and spinal cord which was the first reporting of the pathology of MS. Jean Martin Charcot (1868) described in detail the lesions and reported on the inflammation and structural changes such as demyelination and axonal destruction, naming the condition 'sclérose en plaques disseminées'. Although, the primary cause of MS is still unknown, it has been established that both genetic and environmental risk factors contribute to disease susceptibility (Witte et al, 2014). The genetic component is suggested by familial aggregation in MS and few genes have been identified consistently to influence disease susceptibility which are located on chromosome 6 and are the HLA-DR15 and HLA-DQ6 (Hauser & Oksenberg, 2006). Environmental factors that have been linked with the occurrence of MS are viral and bacterial infections, exposure to vitamin D, smoking and dietary factors (Hauser & Oksenberg, 2006). Nonetheless, there is lack of proof of the pathogenic roles of these factors (Witte et al, 2014).

The most widely used instrument to assess neurological disability in MS population is the Expanded Disability Status Scale (EDSS) and it is considered as a 'gold' standard (Bethoux & Bennett, 2011). It was developed by Kurtzke (1983) to describe disease progression and it consists of an ordinal system ranging from 0 (normal neurological status) to 10 (death due to MS) (Table 2.1). The EDSS is a clinician-administered scale that evaluates the functional systems of the CNS. The measurement of impairments in the lower range of EDSS is based mostly on neurological examination, whilst the upper range (EDSS 4-7.5) depends mostly on the walking abilities of each individual with MS.

Table 2.1 Expanded Disability Status Scale (adapted by Kurtzke, 1983).

Score	Description
1.0	No disability, minimal sings in one FS
1.5	No disability, minimal signs in more than one FS
2.0	Minimal disability in one FS
2.5	Mild disability in one FS or minimal disability in two FS
3.0	Moderate disability in one FS or mild disability in three or four FS. No impairment to walking
3.5	Moderate disability in one FS and more than minimal in several others. No impairment to walking
4.0	Significant disability but self-sufficient and up and about some 12 hours a day. Able to walk without aid or rest for 500m
4.5	Significant disability but up and about much of the day, able to work a full day, may otherwise have some limitation of full activity or require minimal assistance. Able to walk without aid or rest for 300m
5.0	Disability severe enough to impair full daily activities and ability to work a full day without special provisions. Able to walk without aid or rest for 200m
5.5	Disability severe enough to preclude full daily activities. Able to walk without aid or rest for 100m
6.0	Requires a walking aid (e.g. cane, crutch, etc.) to walk about 100m with or without resting

6.5	Requires two walking aids (e.g. pair of canes, crutches, etc.) to walk about 20m without resting
7.0	Unable to walk beyond approximately 5m even with aid. Essentially restricted to wheelchair; though wheels self in standard wheelchair and transfers alone. Up and about in wheelchair some 12 hours a day
7.5	Unable to make more than a few steps. Restricted to wheelchair and may need aid in transferring. Can wheels self but cannot carry on in standard wheelchair for a full day and may require a motorised wheelchair
8.0	Essentially restricted to bed or chair or pushed in wheelchair. May be out of bed itself much of the day. Retains many self-care functions. Generally has effective use of arms
8.5	Essentially restricted to bed much of the day. Has some effective use of arms, retains some self-care functions
9.0	Confined to bed. Can still communicate and eat
9.5	Confined to bed and totally dependent. Unable to communicate effectively or eat/swallow
10.0	Death due to MS

Abbreviations: FS: Functional system

2.2.2 Incidence and prevalence

Multiple sclerosis is the most common cause of chronic neurological disability in early to middle adult life and it has been estimated that the number of pwMS worldwide is 2.3 million (Hauser & Oksenberg, 2006; Browne et al, 2014). Prevalence, has increased from 30 in 2008 to 33 in 2013 per 100,000 (Browne et al, 2014). The pooled prevalence, which is a measure that takes into account the relative size of a country's population, was 1.6 million from the 92 countries that provided data in 2013 (Browne et al, 2014). This increase in prevalence from 2008 to 2013 could possibly be explained by the increased survival rate both in MS and the general population, but it could also be attributed to improvements in diagnosis and reporting of MS (Browne et al, 2014). The incidence of MS was reported in 52 countries, with higher incidence rates been in San Marino (13.75 per 100,000/year), Canada (13.4 per 100,000/year) and Northern European countries (10-12 per 100,000/year) and lower incidence rates in South American countries (0.25-0.6 per 100,000/year) and Eastern Asian countries (0.5-0.63 per 100,000/year) (Browne et al, 2014). It has been estimated that overall there are twice as many women as men living with MS and even though people can be diagnosed at any age, the average age of MS onset is 30 years old (Browne et al, 2014). In UK the prevalence of MS in 2010 was 203.4 per 100,000 population and the incidence was 9.64 per 100,000/year (Mackenzie et al, 2014). The Scottish MS Register that was published in 2018 reported that 425 people diagnosed with MS in 2017 and the incidence for Scotland was 8.6 per 100,000/year, with the higher rate observed in NHS Orkney (17.3 per 100,000/year) and the lowest in NHS Borders (6.2 per 100,000/year) (Scottish MS Register, viewed online 2018). The annual age specific incidence in Scotland

per 100,000 was twice as high for female at the age range 35-39 with a rate of 26.9 as for male at the age range of 40-44 with a rate of 12.1 for the period 2010-2017 (Scottish MS Register, viewed online 2018).

2.2.3 Types of MS

The clinical course of MS may be variable over time and the typical manifestation includes either gradual deterioration of neurologic function, acute episodes of worsening or combination of both. Depending on the current medical status and history, the MS phenotypes can be categorized to either relapsing or progressive. Nevertheless, since MS is a progressive disease the subtype from the initial assessment might change over time, for example from a relapsing-remitting (RR) subtype to evolve to a secondary progressive (SP) subtype (Lublin et al, 2014).

Clinically isolated syndrome (CIS) is defined as the first clinical manifestation of a disease that demonstrates inflammatory demyelination that could be MS, but still needs to fulfil the criteria in time (Miller et al, 2005). The clinical presentation of CIS and early MS has been reported to affect the prognosis over time. Especially, the number of lesions on the magnetic resonance imaging (MRI) can have an effect on the course of the disease and those with an abnormal scan are on a greater risk of having a second diagnosis-defining episode (Morrissey et al, 1993; Jacobs et al, 1997).

The consensus of the definition of RR MS is: “clearly defined disease relapses with full recovery or with sequelae and residual deficit upon recovery; periods between disease relapses characterized by a lack of disease progression” (Lublin et al, 1996) (Figure 2.1). Nearly 85% of MS patients are diagnosed with RRMS

and it is almost twice more likely for females to be diagnosed than males. The clinical course of RRMS can last from years to decades and the relapse rate is highly variable from patient to patient, with average of one or two episodes per year (Trapp & Nave, 2008).

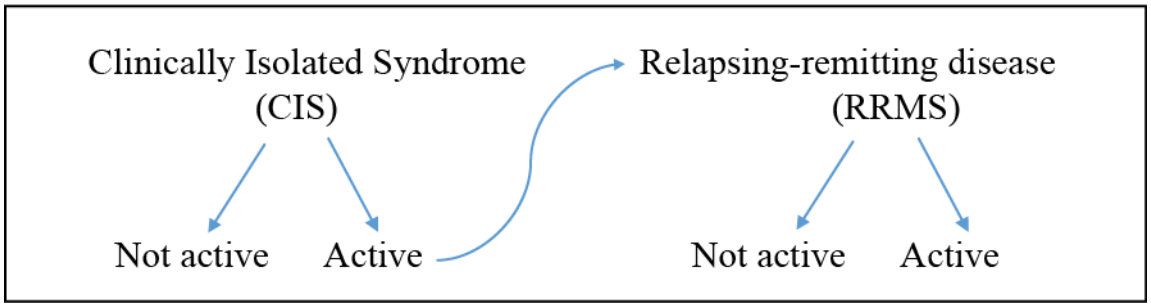


Figure 2.1 Multiple sclerosis phenotype description for relapsing disease.

The definition of primary progressive (PP) MS is as follows: "disease progression from onset with occasional plateaus and temporary minor improvements allowed" (Lublin et al, 1996) (Figure 2.2). Approximately, 15% of pwMS are diagnosed with PPMS and the incidence is similar for females and males, unlike RRMS. The course of the disease begins later in life compared to RRMS and usually relapses are rare or non-existent (Trapp & Nave, 2008).

Secondary progressive (SP) MS is defined as "initial RR disease course followed by progression with or without occasional relapses, minor remissions and plateaus" (Lublin et al, 1996) (Figure 2.2). The majority of RRMS patients typically after eight to twenty years will enter the next phase of the disease. In SPMS stage, there is a continuous and irreversible neurological deficit which is

not related with the relapses and usually affects quality of life, physical and cognitive function (Trapp & Nave, 2008).

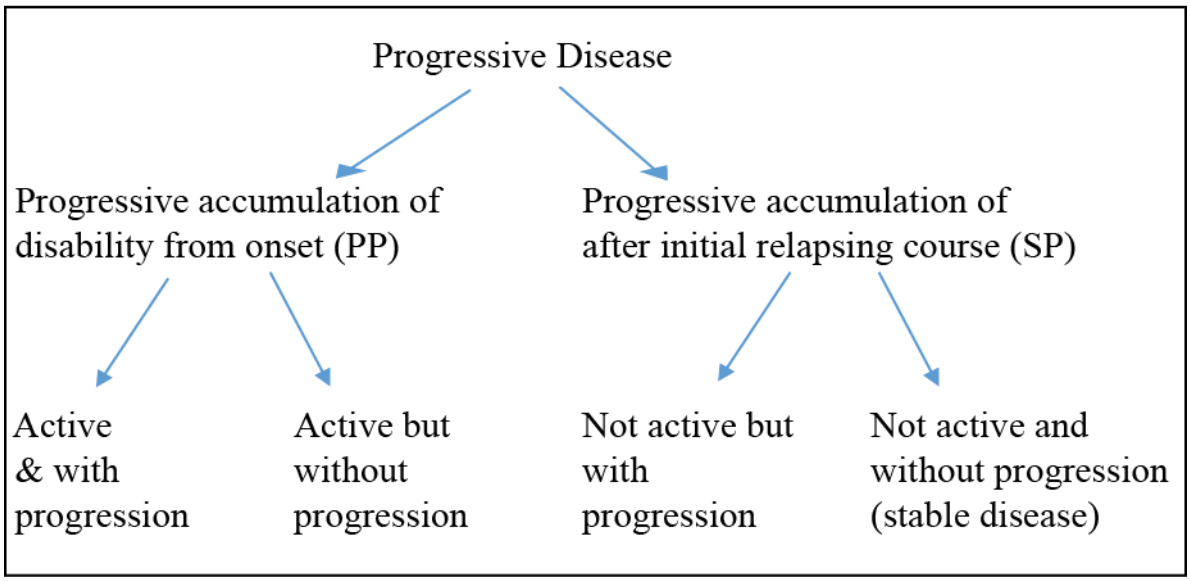


Figure 2.2 Multiple sclerosis phenotype description for progressive disease.

2.2.4 Symptoms

Multiple sclerosis is a chronic demyelinating disease of the CNS with its course being highly variable with many people developing irreversible disability. The initial presentation of MS depends on the location of lesions and the type of onset (i.e. relapsing or progressive) (Brownlee et al, 2017). Along the course of the disease the symptoms that are frequently reported are fatigue, pain, gait impairments, cognitive impairments, spasticity, bladder dysfunction and heat sensitivity and these symptoms are discussed in the section below. Gait impairments will be discussed in more detail in section 2.3.1.

2.2.4.1 Fatigue

Fatigue is a typical symptom of neurological diseases such as multiple sclerosis, post-poliomyelitis, post-stroke and chronic fatigue syndrome (Chaudhuri & Behan, 2004). In the MS population the prevalence has been reported to be 68.7%-78%, with a higher percentage of people with progressive MS (81.1%) reporting fatigue compared to non-progressive types of MS (64.1%). Fifty-six percent of pwMS report that fatigue is the most disabling symptom affecting their ability to perform activities of daily living (Bakshi, 2003; Rooney et al, 2019). A recent systematic review evaluating the pathophysiological pathways of MS fatigue suggested that fatigue mainly originates from a dysfunction of the CNS neural circuits such as greater excitability, less inhibition, reduced cortico-subcortical interaction, secondary to demyelination, increased inflammation and axonal lesions (Langeskov-Christensen et al, 2017). These pathophysiological pathways have been linked to secondary sources of MS fatigue (i.e. cognitive impairments, sleep disturbances and deconditioning) (Langeskov-Christensen et al, 2017).

Despite the impact of fatigue as a disabling symptom in MS population, there are no standard definitions of fatigue. Kluger et al (2013) proposed a unified taxonomy by differentiating perceptions of fatigue that are 'subjective sensations of weariness and increasing sense of effort' from fatigability defined as 'exercise-induced reduction in the ability of muscles to produce force or power whether a task can be sustained'. Measurements of perceived fatigue usually are determined through self-reported questionnaires (i.e. Fatigue Severity Scale, Modified Fatigue Impact Scale, Fatigue Scale for Motor and Cognitive Functions, etc.)

which can be influenced by other symptoms of MS and are completely subjective (Rudroff et al, 2016). Objective quantification of fatigability is usually measured during motor tasks and is characterized by a decline in peak force or power, deceleration of walking speed during endurance walking tests such as the 6-minute walk test (6MWT) or changes in the kinematic or spatiotemporal parameters during gait (Phan-Ba, et al, 2012; McLoughlin et al, 2016; Rudroff et al, 2016; van der Linden et al, 2018). The relation between fatigue and fatigability is unclear. Some studies failed to show a clear association of self-reported fatigue with fatigability which was measured either through recording muscle strength of the hand or by maximal voluntary contractions of the thumb (Iriarti & de Castro, 1998; Romani et al, 2004). However, other studies have found strong associations of perceived fatigue and the decline in force during a sustained maximal contraction either of the hand or lower limb muscles (Andreasen et al, 2009; Steens et al, 2012).

Fatigability is important to consider in the MS population since it affects the ability to perform sustained activities in daily life (Rudroff et al, 2016). Thus far, in a number of studies quantifying fatigability it is suggested that the reduced performance over time is partly of central origin and that it could be a compensatory mechanism for the effects of demyelination on conduction in motor pathways (Ng et al, 2004; Thickbroom et al, 2008; Andreasen et al, 2009). Some studies have examined fatigability as the reduced performance e.g. walking speed during or after longer distance walking tests. People with MS with higher disability levels demonstrated greater deterioration either in walking speed or

kinematic and kinetic parameters during or after the fatiguing tasks compared to mildly impaired pwMS (Phan-Ba et al, 2012; Leone et al, 2016).

In a recent systematic review, Severijns et al. (2017) presented the assessment protocols and outcome measures of fatigability and concluded that currently, no existing protocol can be recommended for clinical use. This is because no studies have compared different protocols to record fatigability and there is limited evidence for the psychometric properties of fatigability outcome measures used so far. Overall, there is a need for standardization among existing protocols assessing fatigability and more importantly to define from the beginning which aspect of fatigability is assessed.

2.2.4.2 Pain

People with MS experience pain and it can be generally categorised as neuropathic, somatic or psychogenic as well as acute or chronic (Crayton & Rossman, 2006). The prevalence of pain in a sample of 1672 patients with MS was 18.1% for dysesthetic pain, 16.4% back pain, painful tonic spasms 11%, Lhermitte's sign 9% and trigeminal neuralgia 2% (Solaro et al, 2004). It has also been reported that different types of pain experienced by pwMS (i.e. back pain, tonic spasms, etc.) are significantly correlated with age, EDSS level, disease type and disease duration (Solaro et al, 2004). It has been demonstrated that pwMS with pain have greater disability (i.e. higher EDSS), more severe symptoms of depression and decreased self-reported overall health compared to pwMS with no pain (Ehde et al, 2003). Since pain in MS can be related to various causes (e.g. diseases course, spasticity, etc.), managing the symptoms may alleviate pain

along with analgesic and anticonvulsant medications if additional relief is needed (Crayton & Rossman, 2006). However, a recent systematic review concluded that there is limited evidence for non-pharmacological interventions (i.e. transcutaneous electrical nerve stimulation, hydrotherapy, transcranial direct stimulation, etc.) and this was due to low methodological quality within the studies (Amatya et al, 2018).

2.2.4.3 Bladder dysfunction

Bladder dysfunction is a common problem in MS and depending on the severity of the neurological disability the report of lower urinary tract symptoms can reach up to 75% and usually is rare at onset of the disease (Marrie et al, 2007; Panicker & Fowler, 2015). The quality of life of those pwMS with bladder symptoms is negatively affected compared to pwMS without any bladder problems and it can also be a psychosocial burden to the patients, the carers but also increase the healthcare costs (Panicker & Fowler, 2015). Previous studies have demonstrated that there is a correlation between bladder and sexual dysfunction in the MS population, since sexual response originates from the nerves on the S2, S3 and S4 spinal levels where bladder innervation also occurs (Betts et al, 1994; Nortvedt et al, 2001). A recent review summarised all the symptoms of the lower urinary tract in MS that comprise by storage symptoms (e.g. overactive bladder, urinary incontinence, nocturia, etc.) and voiding symptoms (e.g. urinary retention, poor stream, etc.) which can affect the quality of life of pwMS (Sakakibara, 2019). Non-pharmacologic interventions for the management of bladder symptoms include pelvic floor exercises, timed voiding, good fluid

intake, but for more severe problems as the course of the disease progresses indwelling catheterization or surgical procedures are offered as treatment options in order to improve the quality of life of pwMS (Gibson & Frank, 2002; Crayton & Rossman, 2006).

2.2.4.4 Heat sensitivity

Over the past century, an ophthalmologist named Wilhelm Uhthoff reported that with increased core body temperature there was a worsening of vision and other MS-related symptoms. This increase of MS-related symptoms with increasing body temperature is called Uhthoff phenomenon (Fraser et al, 2012). It has been hypothesized that this phenomenon is a result of damage to the myelin sheath which results in axonal conduction block when core body temperature is increased (Raminsky, 1973). Heat sensitivity is a common symptom in MS population and Flensner et al. (2011) observed that at least 58% of pwMS report intolerance to heat. Heat sensitivity has been shown to be a significant correlate of many symptoms in the MS population. More specifically, pwMS that are heat sensitive reported higher occurrence of symptoms such as fatigue, weakness in legs, concentration difficulties and pain compared to pwMS that are not affected by sensitivity to increased core body temperature (Flensner et al, 2011). However, evidence for the Uhthoff's phenomenon is non-conclusive. For example, although some studies observed effects of heat sensitivity on cognitive performance (Hämäläinen et al, 2012), other studies reported no significant decrease of cognitive performance in pwMS who had been exposed either to exercise or sauna to increase their core body temperature (Sandroff et al, 2016).

A number of studies have postulated that a relationship exists between heat sensitivity and fatigue. A study comparing resting body temperature in people with RRMS and healthy individuals found that pwMS had elevated body temperature compared to the healthy population and that this raised body temperature was related to fatigue (Sumowski et al, 2014). Similarly, it has been suggested that pwMS who are heat sensitive report higher levels of fatigue and also decreased participation in physical activity compared to pwMS that did not report heat sensitivity (Fjeldstad et al, 2010). Skjerbaek et al. (2013) showed that an increase in core body temperature due to exercise was related to increase in perceived symptoms by pwMS.

Even though Uhthoff's phenomenon may be able to explain the onset of exercise-induced fatigability, the changes in core temperature might not fully explain why fatigability changes over time both during an exercise session and with disease progression. It is of great interest to further investigate the presence of Uhthoff's phenomenon in relation to the gait characteristics in pwMS, in order to understand the underlying mechanisms of exercise-induced fatigability.

2.2.4.5 Depression

Depression is a common symptom in MS with a lifetime rate reaching to 50% and it is related both to the disease itself and the side effects of the disease medications (Sadovnick et al, 1996; Crayton & Rossman, 2006). Previous studies have investigated the clinical impact of depression in pwMS and it was reported that it contributes to decreased cognitive function and quality of life, with increased time away from work and long-term affecting general health with

decreased adherence to treatments (Mohr et al, 1997; Arnett et al, 1999; Wang et al, 2000; Beal et al, 2007). Data from a study on the rate of suicides in the MS population reported that in 15% of all cases the cause of death was suicide and depression was a major risk factor (Sadovnick et al, 1996). Goldman (2005) states that the treatment of depression and depressive symptoms in pwMS should incorporate both psychotherapy and pharmacological treatment (i.e. antidepressants) with a view of the potential adverse effects of certain combinations of medications.

2.2.4.6 Cognitive dysfunction

Cognitive impairment is a common symptom in MS and it may occur at any stage of the disease course. The prevalence in the relapsing-remitting stage and in those with low neurological disability has estimated to be roughly 35%, while in secondary and progressive stages the prevalence of cognitive impairment can reach 60% (Benedict et al, 2006; Patti, 2009). The most commonly affected domains of cognitive function are information processing speed, episodic and working memory that usually reflect damage in specific brain regions (Rovaris & Filippi, 2000; Benedict & Zivadinov, 2011). A recent systematic review summarised the risk factors that could predict cognitive impairment in MS with most noteworthy being the early age of disease onset, progressive disease course, male sex and low average intelligence, as well as certain health behaviours and personality traits, such as smoking (Benedict & Zivadinov, 2011). It has been suggested that cognitive impaired pwMS participate less in social activities and activities of daily living, have a lower working capacity and are more likely to

require personal assistance (Rao et al, 1991; Kalmar et al, 2008). Thus far, even though research into cognitive impairments in MS is increasing, there is inconsistent evidence for any medical or behavioural therapy to improve cognition in MS population (Benedict & Zivadinov, 2011).

2.2.4.7 Spasticity

Spasticity is defined as a 'velocity-dependent increase in tonic stretch reflexes' and often contributes to the disability in MS (Paisley et al, 2002). Previous studies have reported that spasticity affects between 40% and 85% of all pwMS (Paisley et al, 2002; Rizzo et al, 2004). A cross-sectional study highlighted that higher levels of spasticity were significantly associated with disease duration, gait disability and with actively progressing MS (Rizzo et al, 2004). Another important aspect of spasticity is the management and treatment in the MS population. Often, exercise programs including stretching, aerobic and relaxation exercises are used in order to manage spasticity in daily activities (Schapiro & Langer, 1994). However, the most common treatment of spasticity is through a number of medications such as baclofen, gabapentin and botulinum toxin A (Rizzo et al, 2004). Balantrapu et al. (2014) examined the association between leg spasticity and walking performance tests and found that pwMS who presented with leg spasticity (EDSS mean: 6.0) had worse performance in measures such as the 6MWT, timed up and go test (TUG) and the timed 25 foot walk (T25FW), along with decreased walking speed and cadence compared to pwMS without leg spasticity (EDSS mean: 3.5).

2.3 Gait in MS

2.3.1 Gait impairments

As mentioned above, gait impairment is a common symptom in pwMS and this may negatively affect participation and quality of life (QoL) in general. Previous research has established that 85% of pwMS report gait disturbances during the course of the disease (Kelleher et al, 2010; Bethoux & Bennett, 2011) and in some cases in an early stage (e.g. within 15 years of disease onset) (Weinshenker et al, 1989; LaRocca et al, 2011). It has been reported that 50% of pwMS will require assistance for walking within 15 years since diagnosis and that 10% will be wheelchair dependent (Kelleher et al, 2010; Kempen et al, 2016). Whilst there are various factors that may contribute to gait impairment, the major contributors are considered to be sensory changes (proprioception) and resulting poor balance, spasticity, ataxia and lower limb muscular weakness (Cameron & Wagner, 2011).

A considerable amount of literature has been published on the gait patterns of pwMS with different levels of disability. The majority of the studies have examined the gait pattern of pwMS based on the spatiotemporal parameters such as walking speed and step length, with fewer studies also examining joint kinematics in this population. A recent systematic review and meta-analysis aimed to identify and quantify gait deficits in pwMS compared to healthy population. From the 32 studies included in the meta-analysis it was concluded that MS has a significant effect on gait both at a self-selected walking speed and increased pace, even in studies including pwMS with relatively low EDSS. Further, these detrimental effects on walking speed and stride length were found

to be more pronounced in those with higher EDSS levels (Comber et al, 2017). However, this systematic review reported only the spatiotemporal parameters and not the joint kinematics. Further, the range of EDSS of the pwMS included in the studies in the review was relatively broad so conclusions on whether or not people with very low EDSS scores (i.e. less than 2.5) also have gait deficits could not be drawn from this review.

The general consensus is that the typical gait pattern of pwMS is to walk more slowly, with associated shorter stride and step length and prolonged double support phase (e.g. Benedetti et al, 1999; Givon et al, 2009; van der Linden et al, 2014b). As also concluded in the systematic review by Comber et al. (2017) this is true not only for those whose walking ability is affected (EDSS > 4), but also in those with minimal disability. Usually, in minimally impaired pwMS (EDSS < 3.0) walking ability is not affected in daily life, but through assessment via laboratory based measures subtle changes can be observed. It has been demonstrated that even minimally impaired pwMS (EDSS < 3.0) walk slower, with decreased cadence, decreased step length and stride length, increased stride time and increased double support phase (Yahia et al, 2011; Sosnoff et al, 2012; Kalron et al, 2011; Kalron et al, 2013; Kalron et al, 2014). These gait deficits has been also reported to be evident even in adolescents with MS compared to age-matched healthy population (Kalron et al, 2017). One of the first studies examining the gait characteristics of minimally impaired pwMS (EDSS range: 0-2) found changes in all of the aforementioned spatiotemporal parameters compared to a healthy control group. Kinematic changes revealed increase in hip and knee flexion and decrease of ankle dorsiflexion at initial contact at heel strike

and there was an increase in hip flexion and in ankle dorsiflexion during the swing phase of gait (Benedetti et al, 1999). Other studies also demonstrated a decrease of ankle angle at initial contact in pwMS (EDSS < 2.5) compared to healthy individuals (Benedetti et al, 1999; Martin et al, 2006; Kelleher et al, 2010). However, there is no consensus with regard to the hip and knee kinematics in pwMS, with some studies reporting increased hip and knee flexion at heel strike (Benedetti et al, 1999; Kelleher et al, 2010), whilst other authors reporting no significant differences for hip and knee angles in pwMS compared to healthy individuals (Martin et al, 2006; Huisinga et al, 2013; Nogueira et al, 2013). Galea et al. (2017) investigated gait deterioration over a 12-month period in pwMS with EDSS < 3.0 with relapsing remitting MS. Results indicated that this group of pwMS had a statistically significant decrease in walking speed, increase in double support phase and decrease in cadence, but no significant decrease in step length after a one-year period. Interestingly, these changes were not reflected in the clinical status measured with the EDSS. Only one study has examined differences in the spatiotemporal parameters during jogging and inclined jogging between minimally impaired pwMS (EDSS range 0-3.5) and healthy controls (Kalron et al, 2014). It was found that in both conditions pwMS exhibited slower speed, increased step time duration, decreased step length and increased double support phase and step width when compared to a healthy age-matched control group (Kalron et al, 2014).

It is well established from a variety of studies that the gait pattern in pwMS with higher disability (EDSS > 3.5) is altered and is similar with the less impaired population but with more pronounced changes. It has been shown that pwMS

with high EDSS level walk with decreased gait speed, decreased step length and cadence, increased step time and step width and increased double support phase when compared to healthy individuals (Givon et al, 2009; Kelleher et al, 2010; Gianfrancesco et al, 2011; Burschka et al, 2012; Socie et al, 2013). However, few studies have failed to demonstrate statistically significant decrease in cadence (Morris et al, 2002; Kelleher et al, 2010; Remelius et al, 2012; Nogueira et al, 2013) and increase in double support phase (Morris et al, 2002; Givon et al, 2009; Nogueira et al, 2013) when comparing the MS population with a healthy control group.

Another aspect of gait is gait variability (within-session) mostly reported for spatiotemporal parameters. Previous investigations have shown a significant increase in step time variability and step length variability in pwMS compared to healthy individuals, even at an early stage of the disease with minimal disability (EDSS < 3.5) (Flegel et al, 2012; Sosnoff et al, 2012; Socie et al, 2013). Conversely, Kaipust et al. (2012) did not observe any statistically significant changes in step length and step width variability in pwMS compared to a healthy control group, which could possibly be explained by the small sample size (n=10) and the varied disability spectrum of MS population included in the study (EDSS mean 3.95; EDSS range 1-6). Kalron (2016) reported on gait variability of spatiotemporal parameters in a large group of pwMS (n=381) with a wide range of disability level (EDSS range 0-6.5). In the lower end of the EDSS (0-3.5) there were no significant differences for the gait variability parameters, whereas step time variability was significantly increased for the group with EDSS > 4.0 and step length variability was significantly increased to the group with EDSS > 5.0.

One study examined the variability of kinematics in pwMS with a mean EDSS of 3.1 and it was found that there is no significant ankle joint variability measured on two different occasions, but there was increased hip, knee and ankle variability when comparing the MS population with healthy individuals (Crenshaw et al, 2006).

This section has attempted to provide a brief summary of the literature relating to gait deficits in pwMS with different levels of disability. Even though there is a great amount of research investigating gait abnormalities in pwMS during the different stages of the disease progress, there is a need for longitudinal studies to examine the deterioration in gait provisionally providing an understanding of disease progression.

2.3.2 Gait impairments and fatigability

Fatigability is important to consider in pwMS because it affects the ability to perform sustained activities in daily life and to participate in exercise. A recent systematic review by Severijns et al. (2017) reported on the protocols and outcomes measures used to assess and quantify fatigability in pwMS. It was highlighted that based on the International Classification of Functioning (ICF) body function level, most studies used repeated maximal isometric contraction protocols, while for the ICF activity level (e.g. walking performance) protocols inducing changes in performance, such as over a prolonged walking task, were predominantly used (Severijns et al, 2017). Further, it was concluded that there is no gold standard outcome nor protocol to measure fatigability in MS and

fatigability may be influenced by the disability level and the disease phenotype (Severijns et al, 2017).

When focusing on the ICF activity level, most of the existing research has evaluated fatigability by the decrease in walking distance covered over a prolonged walking task. A study by Phan-Ba et al. (2012) recorded the mean walking speed and the deceleration index in a large group of pwMS (EDSS range 0-6) compared to healthy individuals over a timed 100 meter walk (T100mW) and a timed 500 meter walk (T500mW). It was demonstrated that all subgroups of pwMS had a significantly lower mean walking speed compared to the healthy control group. The deceleration index was lower for the whole group of pwMS (i.e. EDSS < 2.0, 2.5-3.5, > 4.0) compared to the control group but it was not significant (Phan-Ba et al, 2012). The same pattern was observed in studies with pwMS with EDSS ranging from 0-6.5 (i.e. groups were divided based on the disability level) over the duration of a 6MWT with a decline in the walking speed in minute six compared to minute one (Burschka et al, 2012; Leone et al, 2016). Conversely, McLoughlin et al. (2016) reported no significant difference in walking speed before and directly after a 6MWT in pwMS with EDSS ranging from 3-6 and further no significant differences between pwMS compared to healthy individuals throughout the length of the test. Recent studies have been carried out to assess fatigability by the distance covered mostly over a 6MWT. A large cross-sectional multinational study (n=208) in pwMS with moderate disability (EDSS mean: 4.2; range: 0-6.5) investigated the distance walk index (DWI) over the duration of a 6MWT (Leone et al, 2016). It was reported that there was a significant decrease of the DWI from minute two to minute six and

further analysis revealed a significant interaction between groups with different EDSS level and time of the 6MWT on the DWI showing that pwMS had a different pattern of DWI over the duration of the 6MWT, i.e. those with higher EDSS scores had increased DWI compared to those lower EDSS scores (Leone et al, 2016). Likewise, similar studies evaluating fatigability with the decline in walking distance in pwMS with moderate disability (EDSS > 3.5), demonstrated a significant decrease of DWI throughout the course of the 6MWT (Dalgas et al, 2014; Proessl et al, 2018).

Studies evaluating gait kinematics over the duration of the 6MWT or protocols of walking to exhaustion have reported mixed results. In a study with 15 pwMS (EDSS range: 4-6) who were using assistive devices (i.e. FES or AFOs), it was found that there was a significant decline in peak dorsiflexion in swing at the end of the 6MWT compared to the start, but there was no significant difference for the ankle angle at initial contact (van der Linden et al, 2018). Similarly, McLouglin et al. (2016) in a similar population group (EDSS range: 3-6) reported a significant decrease of ankle angle at initial contact at the end of a 6MWT with no significant differences in other phases of the gait for the ankle joint, with these changes potentially leading to trips and falls. Deterioration of gait kinematics has also been reported after a task involving walking until a rate of perceived exertion (RPE) of 17 (‘very hard’) or until 60 minutes had elapsed. The authors used a cut-off value on the Fatigue index Kliniken Schmieder (FKS) (i.e. score is based on three-dimensional acceleration measurements of two leg markers) and reported that just over 70% of the group of pwMS (EDSS range: 1-5.5) exhibited fatigability based on their kinematic changes (Sehle et al, 2014).

Although these studies contribute in several ways to our understanding of fatigability and how it affects the walking ability in pwMS, there is a need for standardized fatiguing protocols in order to emulate activities of daily living and outcomes measures that are reliable to capture this phenomenon. Thus far, to the authors' knowledge, no studies have investigated fatigability-induced gait changes in kinematic parameters in pwMS with EDSS < 3.5, which could potentially be used as a prognostic biomarker if fatigability-induced gait deterioration worsens over time.

2.3.3 Foot drop

Foot drop is a common gait impairment in pwMS and is the lack of dorsiflexion during the swing phase of gait (Figure 2.3). Foot drop can be caused by an increased tone in the plantaflexor muscles, weakness of the dorsiflexor muscles and impaired neural control causing co-contraction of agonist and antagonist muscles (Barret et al, 2009). Weakness of the dorsiflexor muscles can be a direct result of MS; damage to the nerves can slow down or disrupt messages, mainly in the spinal cord, making it more difficult to use this muscle group effectively (Kent-Braun et al, 1994). Foot drop can also be temporary phenomenon because of fatigue of the dorsiflexor muscles after walking a certain distance or exercise (Mount & Dacko, 2006).

Even though foot drop is common in many CNS disorders, the mechanisms contributing to foot drop in MS compared to non-progressive disorders (i.e. stroke) might differ. For example, stroke occurs after a sudden interruption of the blood flow in a specific area in the brain and the ischemic area will determine the

type of the deficits (Friedman, 1990). Whilst, for MS, since it is a neurodegenerative disease, foot drop could be a result of an interruption of neural transmission due to inflammation and demyelination of the myelin sheath and axons in the brain and spinal cord (Compston & Coles, 2008).

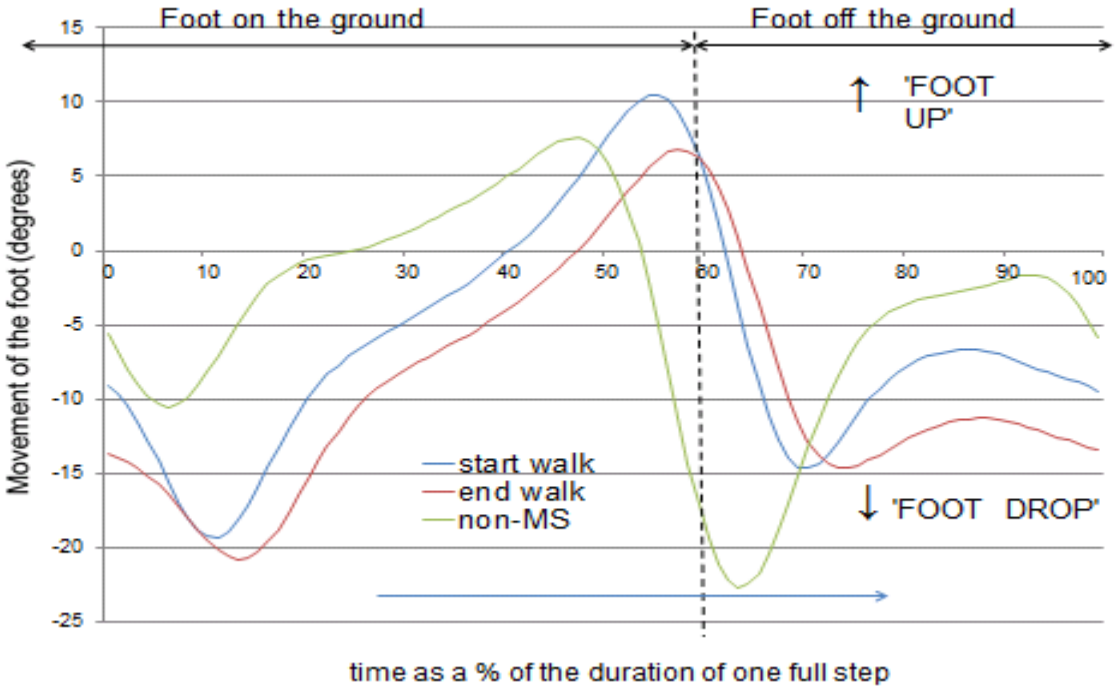


Figure 2.3 The gait cycle of the ankle of a healthy individual (green) and a person with MS (blue: baseline, red: after a 6MWT), with dotted line indicating the toe-off time point of the healthy person.

It has been demonstrated that foot drop can lead to foot dragging, tripping and falling (Gunn et al, 2014). People with MS experiencing foot drop may also have increased fear of falling that could lead to a decline of habitual activity and increased sedentary time, which could have long-term negative implications in their general health (Peterson et al, 2007). In order to compensate for foot drop, whether permanent or temporary, pwMS may adopt strategies such as

circumduction and hip-hiking which may result in an increased effort of walking compared to healthy individuals (Paul et al, 2008).

The prevalence in foot drop in MS population has not been established. However, as part of this PhD, an audit was conducted during the period from September 2016-December 2018, in order to investigate the frequency of *transient* foot drop, activities that initiate transient foot drop and other characteristics such as when it comes on and how long it takes for the transient foot drop to disappear. More details with regard the methods and a table of the full results can be found in Appendix 1. In summary, of the 47 respondents who returned the audit, 70% reported experiencing temporarily foot drop. The majority of people reported that foot drop initiates after walking for an average of 1.6 miles or around 20 minutes, while foot drop could initiate earlier if the people were walking/exercising faster or if they were fatigued. People also reported that foot drop starts after running, cycling or climbing stairs and that they have to stop their activities (n=16) or slow down (n=18). The average time for recovery was reported to be 32 min with a range from one to ninety minutes.

As the disease progresses, foot drop often becomes more established and recovery does not occur anymore. If this occurs, several treatment options exist which aim to maintain function in daily activities and quality of life and reduce the risk of tripping and falling. The most common treatments for foot drop and their impact on the walking ability of pwMS will be discussed in more detail in the next section.

2.3.4 Treatments for foot drop

Conventionally, foot drop has been treated with physiotherapy, such as gait training or strengthening exercises for the dorsiflexor muscles. It has been demonstrated that people presenting foot drop secondary to MS after endurance exercise of the dorsiflexor muscles had an improvement of dorsiflexors control during walking (Mount & Dacko, 2006). The most common way though, is through AFOs and in recent years FES has been gaining popularity.

An AFO is the first line treatment for foot drop and in the UK is funded by the NHS. It can have different forms and comprised by different materials, but always consists of a foot plate and a shin section and can be custom-made for each individual in order to produce the optimal effects (Wening et al, 2013). Existing evidence report that the use of AFOs can reduce the energy cost of walking and can improve static and dynamic balance (Bregman et al, 2012; McLoughlin et al, 2015). However, the use of AFOs has not been shown to significantly increase the walking distance or decreasing the perceived fatigue in pwMS with moderate disability (McLouglin et al, 2015). A qualitative study by Bulley et al. (2015) explored the impact of AFOs on pwMS who presented foot drop. People with MS reported that the use of AFO reduced tripping and falling, provided great stability and balance, but it was enhancing a non-normal gait pattern, was more cumbersome, and discomfort in use (Bulley et al, 2015).

Functional Electrical Stimulation is increasingly used to treat foot drop in neurological disorders such as stroke and MS. Stimulation is applied through surface or implanted electrodes over the common peroneal nerve and the motor point of the tibialis anterior muscle to produce enough dorsiflexion of the ankle

during the swing phase of gait (Burridge et al, 1997). It has been suggested that regular use of FES can strengthen the activation of motor cortical areas and their residual descending connections that might explain possible therapeutic effects of FES on walking speed (Everaert et al, 2010). In two recent qualitative studies, the majority of pwMS reported that the use of FES reduces perceived fatigue and the mental effort of walking, improves the gait pattern, increases confidence and physical activity, but there is a difficulty in the placement of the electrodes with negative financial implications (Bulley et al, 2015; (Miller) Renfrew et al, 2018).

A large and growing body of literature has investigated the effects of FES on walking performance in pwMS. The effects of FES are more commonly described in terms of orthotic and training effects. An orthotic effect is described as the change in walking performance with and without FES. The initial orthotic effect is the change in performance with and without FES the first day of its use, while the continuing orthotic effect is the change in walking with and without FES at a follow up point after a period of regular use of the device. The total orthotic effect represents a combined training and direct effect of the use of FES and is the change in walking with FES at a follow-up point compared to walking without FES at the beginning of the treatment (Taylor et al, 2013). The training effect is the change of walking performance without FES at a follow-up point after regular use of FES relative to the walking performance without FES at the beginning and it describes the impact of regular use of FES on walking over time (Taylor et al, 2013).

The greater part of the literature investigating the orthotic and training effects of FES for foot drop in MS has focused on changes in spatiotemporal parameters of

gait and primarily walking speed over short or long distance walking tests, with fewer studies examining changes in gait kinematics. Many studies have investigated the initial orthotic effect of FES and reported that there is a positive effect on walking speed (i.e. increased walking speed with FES on compared to without FES) mostly over short distance walking tests such as the 10-meter walk test (10mWT) (Taylor et al, 1999; Paul et al, 2008; Barrett & Taylor, 2010; Scott et al, 2013; Downing et al, 2014). Miller et al. (2016) investigated the effect of FES in pwMS walking at self-selected walking speeds (SSWS). It was reported that pwMS with SSWS < 0.8m/s had significant improvements in the walking speed and oxygen cost of gait with the use of FES, whilst the group of pwMS with SSWS > 0.8m/s did not show significant differences in walking speed with the use of FES (Miller et al, 2016). There are several possible explanations for this result, such as the higher disability status (EDSS mean: 6.0) and the potential adjustment to the energy demands of FES of the group with SSWS < 0.8m/s who used FES for a longer period of time compared to the pwMS with SSWS > 0.8m/s. A recent meta-analysis on the effects of FES on walking speed also concluded that there is a statistically significant initial orthotic effect of FES in foot drop in pwMS over short distance walking tests, but only small non-significant initial orthotic effects over longer distance walking tests (Miller et al, 2017). A possible explanation of this might be the presence of fatigability over longer distance walking tests and the benefits of FES might not be sufficient to overcome the gait deterioration. Studies examining changes in gait kinematics have reported that there is a significant increase in peak dorsiflexion in swing and

at ankle angle at initial contact with the use of FES compared to walking without FES (Scott et al, 2013; van der Linden et al, 2014a; van der Linden et al, 2014b). The training effect of FES on walking performance in pwMS has also been documented, although in fewer studies and not with consistent findings. A study examining the training effect of FES, indicated improvements in ankle angle at initial contact when walking without FES at the end of 12-week period compared to unassisted walking at baseline (van der Linden et al, 2014a). Stein et al. (2010) investigated the long-term training effect of FES over an 11-month period in both progressive (MS) and non-progressive (stroke) conditions. It was demonstrated that there is an improvement in walking speed over a 10mWT and a 4-minute walk test at 3-month follow-up for both stroke and MS patients, but at the 11-month follow up a training effect was seen only the stroke population (Stein et al, 2010). Contrary to these findings, studies evaluating the training effect of FES for periods ranging from 18-20 weeks have failed to find significant changes in walking speed over short distance walking tests (Barrett et al, 2009; Barrett & Taylor, 2010; Street et al, 2015). Even though Street et al. (2015) did not observe a significant training effect of walking speed after 20 weeks of FES used, 31% of the participants achieved a clinically meaningful training effect (> 0.05m/s). Training effects have been consistently found in non-progressive conditions, such as stroke (Kafri & Laufer, 2014). However, this has not been the case for MS, which can partly be explained by the inflammatory and degenerative course of the disease.

A number of authors have also shown that the use of FES can have a beneficial effect on the quality of life and the perceived walking ability of pwMS, with

improvements in activities of daily living and reducing the number of falls (Barrett & Taylor, 2010; Esnouf et al, 2010; Mayer et al, 2015). These perceived benefits of FES for foot drop have the potential to increase social participation, community mobility and daily activity for these individuals. One of the perceived benefits that pwMS report is the reduced concentration and mental effort during walking (Bulley, et al, 2015; (Miller) Renfrew et al, 2018). Thus far, these perceived benefits of FES have not been quantified and the aim of Chapter 6 is to explore the benefits of FES under dual-task and fatiguing conditions that pwMS experience daily.

2.4 Psychometric properties and outcome measures

2.4.1 Psychometrics of outcome measures

Psychometric properties of outcome measures are essential information in both research and clinical practice. Outcome measures are utilized to provide information on a patient's status, which can predict the success or failure of a treatment or an intervention based on the scores obtained (Mokkink et al, 2010a). Therefore, outcome measures need to have robust psychometric properties, such as reliability, validity and responsiveness, to allow for correct interpretation of the results. However, there is a lack of consensus about the definition and terminology used to describe measurement properties, along with how these measurement properties should be assessed (Terwee et al, 2012). The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) initiative was developed in order to reach consensus on how these measurement properties should be defined (Table 2.2) and to develop

standards of how these measurement properties should be evaluated in terms of study design and statistical analysis (Mokkink et al, 2012). If the methodological quality of a study on measurements properties of an outcome measure is appropriate, the results can be more robust for use of the specific outcome studied, whilst if the methodological quality of a study is inadequate the results cannot be trusted and the quality of the outcome measure remains unclear (Higgins & Green, 2008). The COSMIN taxonomy proposes three main domains, which are reliability, validity and responsiveness and each domain contains one or more measurement properties (Mokkink et al, 2012).

2.4.1.1 Reliability

The overall definition of the reliability domain is the extent to which the scores of patients that have not changed are the same for repeated measurements and the degree to which the measurement is free from measurement error (Mokkink et al, 2010b). The reliability domain is sub-divided in three measurement properties, which are internal consistency, reliability and measurement error. Reliability and measurement error will now be considered in more detail given the focus of Chapter 4.

The reproducibility of the results when a test is assessed in a repeated measures design over time is known as test-retest reliability; inter-rater reliability is when a test is implemented by different persons on the same occasion and intra-rater reliability is assessed when comparing the results from the same rater on different occasions (Mokkink et al, 2010b).

The consistency of the results is considered the degree to which the results do not change over repeated measurements, whereas agreement refers to how close the

results of repeated measurements are, by estimating the measurement error over repeated measurements (de Vet et al, 2006b).

Table 2.2 Definitions of domains and measurement properties adapted by Mokkink et al. (2010b).

Domain	Measurement property	Definition
Reliability		The degree to which the measurement is free from measurement error: The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions, e.g. using different sets of items from the same health related patient-reported outcomes (HR-PROs) (internal consistency), over time (test-retest) by different persons on the same occasion (inter-rater) or by the same persons on different occasions (intra-rater)
	Internal consistency	The degree of interrelatedness among the items
	Reliability	The proportion of the total variance in the measurements which is because of 'true' differences among patient
	Measurement error	The systematic and random error of a patient's score that is not attributed to the true change in the construct to be measured
Validity		The degree to which an instrument measures the construct(s) it purports to measure
	Content	The degree to which the content of an instrument is an adequate reflection of the construct to be measured

Domain	Measurement property	Definition
	Construct	The degree to which the scores of an instrument are consistent with hypotheses (i.e. with regard to internal relationships, relationships to scores of other instruments or differences between relevant groups) based on the assumption that the instrument validly measures the construct to be measured
	Criterion	The degree to which the scores of an instrument are an adequate reflection of a 'gold standard'
Responsiveness		The ability of an instrument to detect change over time in the construct to be measured

Intra-Class correlation coefficients (ICC) are the most common method used to report reliability of an outcome measure in terms of its consistency. The ICC value obtained, between -1 and +1, demonstrate the strength of the relationship but also the direction. Zero values indicate no relationship, with ICC values of ≥ 0.75 regarded as excellent level of practical and clinical significance for reliability, while 'good' was between 0.60-0.74 and 'fair' between 0.40-0.59 (Shrout & Fleiss, 1979; Cicchetti, 1994). Several ICC models exist and it is important to choose the correct one depending on the number of raters or tests included in the analysis (Shrout & Fleiss, 1979; Koo & Li, 2016). Two other methods of analysis are sometimes used to report consistency that are the Pearson correlation and Kappa coefficient (Streiner et al, 2014). Pearson's correlations are used to assess the relationship between two sets of results. It has been suggested that Pearson coefficient is inadequate to use as it measures the linear relationship and not systematic differences (Streiner & Norman, 2008). Kappa coefficient is most appropriate when only two levels of outcome are expected; for example, when something is either present or absent in a test result (Streiner et al, 2014). Reliability parameters are highly dependent on the variation of the sample population and can be generalizable only to populations with similar variation (de Vet et al, 2006b).

Measurement error (i.e. absolute agreement) is any systematic and random errors that are not attributed to true change in the construct to be measured (Mokkink et al, 2010b). Agreement parameters are a characteristic of the instrument itself and are expressed on the actual scale of the instrument (de Vet et al, 2006b). The indices that are widely used to assess measurement error are the standard error of measurement (SEM), minimal detectable change (MDC) or smallest detectable change (SDC) and the limits of agreement

(LoA) obtained through the Bland and Altman plots. The SEM represents the variance between subjects or measurements and equals to the square root of the error of variance. The clinical interpretation of the SEM can be expressed as the MDC, as it provides information of the cut-off point above which a change can be regarded as 'true' change (de Vet et al, 2006b). The Bland & Altman plot was developed to describe the agreement between two different instruments or the repeatability of one instrument on repeated occasions. The graph is produced by plotting the mean difference of each of the participants data against the difference in their scores on the same two occasions and any systematic differences or obvious outliers can be seen graphically (Bland & Altman, 1999).

2.4.1.2 Validity

According to the COSMIN taxonomy, the overall definition of validity is the degree to which an instrument measures the construct(s) it purports to measure. The three sub-divisions in this domain address slightly different aspects of validity (Mokkink et al, 2010b). Content validity, which is relevant mostly for questionnaires, is the degree to which the content of an instrument is an adequate reflection of the construct to be measured (Guyatt et al, 1993). The degree to which the scores of an instrument are consistent with the hypotheses (i.e. with regard to internal relationships, relationships to scores with other instruments or differences between relevant groups) is defined as construct validity. Criterion validity is the degree to which the scores of an instrument are an adequate reflection of a gold standard (Mokkink et al, 2010b). The most common method of analysis for validity is either the Pearson correlation or the Spearman's rank

correlation coefficient. The strength of correlations is regarded as small if the values are between 0.1-0.3, moderate if they are between 0.3-0.5 and strong between 0.5-1.0 (Portney & Watkins, 2000).

2.4.1.3 Responsiveness

The term responsiveness is defined as the ability of an instrument to detect clinically important changes over time, even if these changes are small (Guyatt et al, 1989). It has been considered by some to be a component of longitudinal validity (Terwee et al, 2003). However, Kirschner & Guyatt (1985) specified that an outcome measure has three properties; firstly reliability which characterizes intra-subject variability, secondly validity which outlines that any changes detected should be consistent with an external standard and finally responsiveness which should detect clinically important changes. There are several indices that measure the magnitude of change, such as effect size (ES), standardised response mean (SRM), paired t-tests, the Guyatt Responsiveness Index (GRI) and the Receiver operating characteristics (ROC) curve (Terwee et al, 2007). Another measure for responsiveness is the minimal clinically important difference (MCID) that provides information of how much change is enough change in an outcome measure to detect change in a clinical condition and focuses at an individual level (Terwee et al, 2007). Sensitivity to change to a treatment is seen as a characteristic of both the treatment and the variance of the population tested, and therefore responsiveness of an outcome measure is distinct for each clinical population (Beckerman et al, 2001).

2.4.2 Outcome measures of walking performance in MS and psychometric properties

As mentioned in a previous part of this literature review, deterioration of gait is a common symptom in pwMS as it affects mobility and quality of life. Therefore, measurement of gait is a critical part of a patient's assessment, as it can provide information into the clinical disease status and assess the efficacy of symptomatic and rehabilitation therapies. The outcome measures used to assess walking performance need to demonstrate robust psychometric properties, in order to allow meaningful interpretation of the results. Further, the feasibility of administering walking assessments (i.e. equipment and personnel needed), space and time requirements of both personnel and patients need to be regarded (Bethoux & Bennett, 2011). Particularly, in pwMS the interpretation of changes in scores of walking performance outcomes may be challenging due to the nature of the disease with fluctuations in performance from day to day, or time of the day being assessed and the presence of other impairments along with fatigue (Albrecht et al, 2001; Morris et al, 2002; Crenshaw et al, 2006). As a result of this intrinsic variability the reliability of a walking performance outcome might be affected, and therefore studies examining reliability of an outcome measure should consider taking into account the disability or walking impairments level of the MS population.

Various outcome measures have been utilized to evaluate walking performance in the MS population, by either self-reported questionnaires and rating scales or quantitative measures of gait characteristics such as motion analysis. Before the start of this PhD, two narrative reviews had summarised the outcome measures that have been used to assess walking performance in the MS population and provide more insight into their psychometric properties (Bethoux & Bennett, 2011; Kieseier & Pozzilli, 2012).

2.4.2.1 Clinical rating scales and self-reported outcomes of walking performance

The EDSS in MS is the most common instrument to characterize disability and disease progression (Kurtzke, 1983). Walking ability is assessed in the middle range of the scale (4-7.5) and is based on the maximum distance walked by a patient. The characterization of disability with EDSS must be interpreted cautiously in this range, since there is a wide variability in the walking distance due to day to day fluctuations in a patient's performance (Albrecht et al, 2001) or in environmental factors that can affect the distance walked. Interestingly, there is limited evidence for the psychometric properties of EDSS, demonstrating validity but generally poor reliability and responsiveness to change (Sharrack et al, 1999; Hobart et al, 2000). Even though EDSS is a useful instrument to assess disease severity in MS, it should probably not be used as an outcome to measure walking ability in routine clinical assessment (Bethoux & Bennett, 2011).

The Hauser Ambulation Index (HAI) converts ambulation-related disability based on the T25FW into an ordinal scale. It is a 10-point scale (0-9), with 0 representing no impairment and 9 representing confinement to wheelchair and inability to transfer independently (Hauser et al, 1983). It has been demonstrated that the HAI exhibits excellent reliability, but it has shown weak responsiveness to clinical change (Sharrack et al, 1999). The HAI is useful tool to classify pwMS based on their walking performance and it is easier to administer than EDSS, although the weak responsiveness makes it less suitable for measuring performance after interventions (Bethoux & Bennett, 2011).

The 12-Item MS Walking Scale (MSWS-12) was originally developed as patient-reported based outcome to capture the impact of MS on walking ability. It contains 12 items with

Likert-type responses with a recall period of two weeks. The higher the score, the greater the impact on walking ability (Hobart et al, 2003). The psychometric properties have been extensively examined in diverse MS population and in both community and hospital settings. The MSWS-12 has been shown to have strong internal consistency, excellent reliability, validity, and good generalizability and responsiveness to change (Hobart et al, 2003; McGuigan et al, 2004; Motl & Snook, 2008; Baert et al, 2014; Learmonth et al, 2013a; Andreopoulou et al, 2018). Comparing the MSWS-12 with other measures of physical and cognitive domains, showed stronger correlations with measures of physical domain that relate to mobility and lower extremity function (Motl et al, 2008). A strong correlation was found between the MSWS-12 and the EDSS between 1 to 4.5, whilst a weak correlation with EDSS scores ranging from 5 to 8. A strong correlation has been shown also with accelerometer counts and the MSWS-12, suggesting that there is a relationship between perceived and objective walking ability (Motl et al, 2008).

The Rivermead Visual Gait Assessment (RVGA) is a measure of gait performance that was developed for clinical use in patients with neurologic diseases. It comprises of 20 items that are scored in a 4-point scale (0: normal; 3:severe) depended on the joint movements during the swing and stance phase of gait (Lord et al, 1998). The RVGA has demonstrated good inter-rater and intra-rater reliability and sensitivity to treatment effects (Lord et al, 1998).

2.4.2.2 Objective timed walking tests

Objective walking tests can provide a quantitative measure of walking performance, either by recording the walking speed or the walking distance covered. There are short- and long-

distance walking tests and even though they are not disease-specific they have been extensively utilized in the MS population.

The T25FW was initially seen as part of the HAI (Hauser et al, 1983) and was later integrated as part of the Multiple Sclerosis Functional Composite (MSFC) along with the nine-hole peg test (9-HPT) and the paced auditory serial addition test (PASAT) (Fischer et al, 1999). The person is asked to walk as fast and safely as he/she can across a line of 25ft. course without turns and the time to complete the task is recorded. Even though it was not designed as a disease-specific outcome for the MS population, it has been used widely in clinical and research settings and has been extensively examined for its psychometric properties. It has been found to have excellent test-retest, inter-rater and intra-rater reliability, with no apparent practice effects (Rosti-Otajarvi et al, 2008; Learmonth et al, 2012; Learmonth et al, 2013a; Hobart et al, 2013). The T25FW has been reported to have strong correlations with the MSWS-12, the 6MWT and the 100m walk test (Goldman et al, 2008; Cavanaugh et al, 2011; Phan-Ba et al, 2011; Hobart et al, 2013). The responsiveness of the T25FW has been examined after interventions, such as intravenous methylprednisolone (IVMP), dalfampridine or fampridine treatments and physical rehabilitation, and has been able to capture clinically important changes of the walking ability (van Winsen et al, 2010; Filipovic et al, 2011; Coleman et al, 2012; Baert et al, 2014; Jensen et al, 2016). There is a broad acceptance that a 20% change in the time taken to complete the T25FW is a meaningful change in walking performance in the MS population (Kragt et al, 2006; Coleman et al, 2012; Hobart et al, 2013; Cohen et al, 2014). Another short distance walking test that has been widely used in the MS population to assess the walking performance is the 10mWT. The person is instructed to walk either at

a self-selected or at a fastest speed in a 10m course. The inter-rater and test-retest reliability have been found to be high for the 10mWT both at normal and fastest walking speeds (Vaney et al, 1996; Nilsagard et al, 2007; Paltamaa et al, 2005; Feys et al, 2014). It was reported that there is a strong correlation between the 10mWT and the 6MWT at both normal and fast speeds (Gijbels et al, 2012). Strong correlations were observed with measures of participation such as the Frenchay Activities Index (FAI) and accelerometer counts (Kierkegaard et al, 2011; Stellman et al, 2015), but only weak correlations were reported with self-reported fatigue (Morris et al, 2002). The 10mWT has demonstrated adequate responsiveness in both mildly disabled pwMS and in patients with EDSS range 0 to 6 (Nilsagard et al, 2007; de Groot et al, 2006). However, many studies that have examined responsiveness of the 10mWT did not apply any interventions in the interim period, in order to record clinically meaningful changes (Paltamaa et al, 2008; Kempen et al, 2011; Freeman et al, 2013). These short distance walking tests (i.e. 10mWT, 30mWT, etc.) have various starting instructions (static vs dynamic) and pacing rhythms (normal vs fast). The differences in administration can affect the results of the assessments and therefore a more unified protocol would be useful to compare results among studies (Graham et al, 2008).

Longer distance walking tests are commonly record the total distance walked as a measure of walking performance. The most widely used test is the 6MWT in which the person is asked to walk for six minutes at a maximal speed either in straight, square or ellipse course which involves turns (Goldman et al, 2008). The 6MWT has demonstrated excellent test-retest and inter-rater reliability (Paltamaa et al, 2005; Fry et al, 2006; Goldman et al, 2008; Feys et al, 2014), with an MDC of 76.2m for pwMS with EDSS range 5-6.5 (Learmonth

et al, 2012) and an MDC of 88m for pwMS with mild to moderate disability (Learmonth et al, 2013a). It has been shown to have strong correlations with EDSS level and the MSFC and with self-reported measures such as the MSWS-12 (Goldman et al, 2008). Several studies have examined the responsiveness of the 6MWT and it was found that it is a sensitive tool to capture clinically meaningful changes over time (Paltamaa et al, 2008; Baert et al, 2013; Freeman et al, 2013). However, pwMS that experience severe fatigue and with moderate or severe disability levels, walking for six minutes can be exhausting and they might have to rest during the test. It has been suggested that the 6MWT is most likely a measure of walking endurance (Bethoux & Bennett, 2011). An alternative is the 2-minute walk test (2MWT), in which the same protocol applies by walking at maximal speed for two minutes and the distance covered is recorded (Gijbels et al, 2011). Its psychometric properties have not been extensively documented in MS population, but few studies that have examined them reported good reliability and responsiveness to change after IVMP therapy or physical rehabilitation interventions but no MICD values have been reported (Filipovic et al, 2011; Baert et al, 2013; Feys et al, 2014).

2.4.2.3 Quantitative gait analysis

Walking performance tests are widely used and are simple in administration in clinical and research settings, but their main drawback is that they detect only deviations from normal walking (i.e. decreased walking speed or walking distance) and its variation over time without providing information on the gait pattern and thus possible underlying mechanisms (Bethoux & Bennett, 2011).

Three-dimensional gait analysis (3DGA) through motion capture systems has been considered the 'gold' standard in terms of quantitative gait analysis (Bethoux & Bennett, 2011; Cofré Lizama et al, 2016) and it is one of the most common outcome measures used in MS population (Andreopoulou et al, 2018). The 3DGA utilizes camera systems to track marker trajectories of passive (e.g. Vicon, Oxford, UK) or active infrared emitting diode (e.g. Optotrak, NDI, Waterloo, Canada) markers that are placed on anatomical landmarks of the lower limbs. Many studies have reported gait deterioration based on kinematics changes even in minimally affected pwMS at an early disease stage (Martin et al, 2006; Sosnoff et al, 2012; Galea et al, 2017). However, the space and time requirements, cost of equipment and the complexity of data analysis have been reported as drawbacks for its use in a clinical setting for routine gait assessments (Bethoux & Bennett, 2011). A recent topical review on the use of gait analysis in pwMS suggested that gait measures obtained by 3DGA could potentially be used as sensitive biomarkers not only of mobility, but also of disease progression since they allow a better understanding of underlying mechanisms of walking disability that cannot be detected by conventional spatiotemporal parameters of gait (Cofré Lizama et al, 2016). However, evidence of the psychometric properties of 3DGA in MS population is lacking (Andreopoulou et al, 2018).

Another outcome measure that have been used in the MS population to quantify gait impairments is the GAITRite (CIR System, Inc, Havertown, PA). It consists of an instrumented walkway with sensors to identify footfall contacts, enabling the quantification of spatiotemporal parameters of gait (Cutlip et al, 2000). Spatiotemporal parameters measured by the GAITRite have shown a strong correlation with the T25FW and with EDSS (4-6), but only moderate correlation with the MSWS-12 (Sosnoff et al,

2011b). Further investigation of its reliability and responsiveness is needed, in order to potentially be used to capture changes in the context of rehabilitation or research practice.

2.4.2.4 Outcomes of walking performance and measures of participation in daily life

Pedometry and accelerometry have been used to detect movement by either counting of the steps or by monitoring movement in more dimensions. They have been utilized in studies with pwMS to provide an objective assessment of physical activity (Pearson et al, 2004; Gosney et al, 2007; Motl et al, 2007). Motl et al. (2007) reported that the use for seven days of either an accelerometer or pedometer is a reliable measure (ICC=.93) to estimate physical activity in MS population, with a minimum of three days yielding an ICC of 0.80. Moreover, it has been demonstrated that habitual walking performance (i.e. amount of steps in a customary environment based on accelerometry) was significantly predicted by walking tests such as the 2MWT and the 6MWT (Gijbels et al, 2010). Further validation and examination of the responsiveness to changes have not been investigated in MS population (Bethoux & Bennett et al, 2011), which is of importance with regard to day-to-day variability in MS.

There is no 'gold' standard measure to assess walking performance in pwMS and is important that both clinical and patient perspectives are taken into account to evaluate function and disease status. Three-dimensional gait analysis is often considered the 'gold' standard to assess even minimal changes in walking performance and has been widely used in the MS. However, its psychometric properties has not yet been investigated in pwMS. It is of major importance that all outcome measures used to evaluate walking

performance are psychometrically sound in order to provide robust information to both clinician and researchers.

2.4.3 Outcome measures of fatigue and fatigability in MS and psychometric properties

Fatigue is one of the most debilitating symptoms in pwMS, with high prevalence rates as mentioned previously. However, there are challenges in the evaluation of fatigue because of the subjectivity and the multidimensionality of the symptom (Flachenecker et al, 2002). Perceived fatigue is measured through self-reported instruments where people rate or describe the impact or severity of their fatigue. Performance-based fatigability is assessed after a period of sustained level of physical performance (Kluger et al, 2013).

The measurement of fatigability can be valuable for both clinical and research purposes. The majority of the studies evaluating fatigability have used either sustained maximal contractions of the thumb or lower limb muscles (de Haan et al, 2000; Ng et al, 2004; Andreasen et al, 2009), or by performing long distance walking protocols and measuring the performance in walking speed, the deceleration from the start to the end of the task and kinematic changes of the gait pattern (Phan-Ba et al, 2012; Sehle et al, 2014; McLoughlin et al, 2016). Nevertheless, a recent systematic review by Severijns et al. (2017) synthesized all the outcome measures and protocols that have been used so far to measure fatigability in pwMS. This review concluded that there are no standardized protocols to assess and comparing findings among studies assessing fatigability in this population and no psychometric studies have thus far been found (Severijns et al, 2017).

Fatigue, on the other hand, is measured though self-reported questionnaires that have been developed. There are many generic and disease specific fatigue questionnaires and can

provide either a multidimensional or unidimensional assessment of fatigue. These questionnaires might measure different aspects or different theoretical constructs of fatigue, so in order to select the most appropriate a clinician or researcher must consider the underlying concept of fatigue that need to be captured, the psychometric properties of the instrument and the practical feasibility (Dittner et al, 2004; Kos et al, 2004). Table 2.3 presents an overview of all the self-reported fatigue questionnaires that have been examined for their psychometric properties in pwMS and have been used to assess perceived fatigue in this population.

Thus far, only two systematic reviews have evaluated the psychometric properties of self-reported fatigue questionnaires in pwMS (Kos et al, 2004; Elbers et al, 2012). Elbers et al. (2012) reported that 20 questionnaires have been used in MS population and have been examined for their psychometric properties, with the most frequently used questionnaire being the Fatigue Severity Scale (FSS).

One review, suggested the use of Fatigue Impact Scale (FIS) and Modified Fatigue Impact Scale (MFIS) for assessing self-reported fatigue (Kos et al, 2004). However, this study did not examine the methodological quality of the studies assessing the psychometric properties by standards guidelines (e.g. COSMIN) and the recommendations should be interpreted with caution. The other systematic review on psychometric properties recommended the use of the Fatigue Scale for Motor and Cognitive Function (FSMC) since it was found to have moderate level of evidence for internal consistency and structural validity and the Unidimensional Fatigue Impact Scale (U-FIS) that showed adequate reliability and structural validity in the MS population (Elbers et al, 2012).

Clinicians and researchers who wish to assess fatigue should consider whether a particular tool reflects the aspect of fatigue that they are interested in and it has sound psychometric properties. Future studies should consider investigating apart from fatigue and other contributing factors to fatigue such as mood, depression and sleep.

Table 2.3 Description of fatigue questionnaires that have been used in the MS population.

Questionnaire	Construct assessed	Recall period	Description
CFQ	Fatigue severity	Last day	It consists of 11 items measuring the severity of mental and physical fatigue. Each item is scored from 0 (less than usual) to 3 (much more than usual), with higher scores representing high levels of fatigue
CIS-20R	Impact of fatigue/Fatigue severity	Last 2 weeks	Consists of 20 items and responses are based on a 7-point Likert type scale (1-7), with 1 'yes, that it is true' to 7 'no, that is not true'
D-FIS	Impact of fatigue	Last day	It is a unidimensional scale consisting of 8 items; rating is from 0 'no problem at all' to 4 'extreme problem'
EMIF-SEP	Impact of fatigue	Last month	It is the adapted French version of the FIS and is composed by 41 items on 4-point Likert scale (1=it's always false to 4 'it's always true')
FAI	Impact of fatigue/Fatigue severity	Last 2 weeks	It consists of 29 items on a 7-point Likert scale, with lower scores indicating less impact and severity of fatigue
FDS	Fatigue severity	Not specified	This scale consists of 5 items rated in a scale from 0-3; the total score ranges from 0-17 and the higher it is the greater the fatigue

Questionnaire	Construct assessed	Recall period	Description
FIS	Impact of fatigue	Last month	It is a 40-item questionnaire with rating ranging from 0 ‘no problem’ to 4 ‘extreme problem’
FSMC	Impact of fatigue/Fatigue severity	General	It consists of 20 items measuring mental and physical fatigue, with rating from 1 ‘does not apply at all’ to 5 ‘applies completely
FSS	Impact of fatigue/Fatigue severity	Last week	It consists of 9 items with 1 indicating ‘strongly disagree’ to 7 indicating ‘strongly agree’
FSS-7	Impact of fatigue/Fatigue severity	Not specified	It is the same as the FSS after removing item 1 and item 2 from the original version
FSS-5	Impact of fatigue	Not specified	It is the same as the FSS after removing item 1, item 2, item 6 and item 8
MFI	Impact of fatigue	Lately	It is 20-item questionnaire on a 5-point Likert scale, with higher scores indicating greater fatigue
MFIS	Impact of fatigue	Last month	It is a 21-item questionnaire that is divided in physical, cognitive and psychosocial subscales. The scoring ranges from 0 ‘never’ to 4 ‘almost always’

Questionnaire	Construct assessed	Recall period	Description
MFSS	Factors influencing fatigue	Not specified	It is a unidimensional questionnaire consisted of 6 items and is rated on 7-point Likert scale. Higher scores indicate greater fatigue
NFI-MS	Fatigue severity/Factors influencing fatigue	Last 2 weeks	It consists of 33 items which are rated on 4-point Likert scale ranging from 0 ‘strongly disagree’ to 3 ‘strongly agree’
PS-F	Impact of fatigue	Last month	It is a unidimensional scale consisted of 16 items; it is rated on 6-point Likert scale and higher score indicates greater fatigue
SOFI	Fatigue severity	Last 6 months	It consists of 20 items, in which feelings of being tired are rated from 0 ‘not had such feelings at all’ to 6 ‘had such feelings to a very high degree’
U-FIS	Impact of fatigue	Last week	It is 22-item scale that is a modified version of the FIS; it is rated on a 4-point Likert scale with higher scores indicating greater fatigue
WEIMUS	Impact of fatigue	Last 2 weeks	It consists of 17 items rated on a 5-point Likert scale; the scale has cognitive and physical sub-scores, with higher scores indicating higher degree of fatigue

Questionnaire	Construct assessed	Recall period	Description
VAS	Impact of fatigue/Fatigue severity	Not specified	Three VAS to assess impact of fatigue on daily life, self-care activities and household and occupation; the answer line of 10mm ranges from ‘no influence at all’ to ‘a lot of influence’

Abbreviations: CFQ: Chalder Fatigue Scale; CIS-20R: Checklist Individual Strength Questionnaire; D-FIS: Daily Fatigue Impact Scale; EMIF-SEP: French adaptation of Fatigue Impact Scale; FAI: Fatigue Assessment Inventory; FDS: Fatigue Damage Spectrum; FIS: Fatigue Impact Scale; FSMC: Fatigue Scale for Motor and Cognitive Function; FSS: Fatigue Severity Scale; MFI: Multidimensional Fatigue Inventory; MFIS: Modified Fatigue Impact Scale; MFSS: Multiple Sclerosis-Specific Fatigue Severity Scale; NFI-MS: Neurological Fatigue Index; PS-F: Performance Scale Fatigue; SOFI: Swedish Occupational Fatigue Inventory; U-FIS: Unidimensional Fatigue Impact Scale; WEIMUS: Wurzburg Fatigue Inventory for Multiple Sclerosis; VAS: Visual Analogue Scale

2.5 Dual-task performance in MS

As discussed in previous parts of this literature review, one of the benefits of the use of FES reported by pwMS themselves is that it reduces the mental effort of walking and that as a result less concentration is needed on the walking task (Bulley et al, 2015; (Miller) Renfrew et al, 2018). It has been suggested that postural control and cognition compete for a common pool of attentional resources and if one task is becoming more challenging, the available resources reach their limit and the performance in one or both tasks will deteriorate (Stins & Beek, 2012). Another theory, termed as bottleneck theory, assumes that due to limited resources there is a point in information processing that only one task can be performed at a time, causing a decline in the other while dual-tasking (Pashler, 1994). The execution of a motor task simultaneous with a cognitive task is termed dual-tasking and the cognitive-motor interference can be quantified by the dual-task cost (DTC), which is the percentage of change in performance from single to dual task (Yogev-Selingmann et al, 2012).

Cognitive tasks vary depending on demands and mental processes required to execute them. The majority of cognitive tasks that have been utilized in studies with MS population can be categorized in four domains. The definitions of these domains are based on a previous meta-analysis and include mental tracking, verbal fluency, discrimination and decision-making and working memory tasks (Al-Yahya et al, 2011). Mental tracking tasks are used to assess sustained attention and information processing, since they require to hold information in the mind while performing a mental manipulation process (e.g. serial digit subtraction and naming alternate letters of the alphabet) (Al-Yahya et al, 2011).

Verbal fluency tasks examine executive function and require word production spontaneously and under pre-specified conditions with a common example been the word list generation (Al-Yahya et al, 2011). Discrimination and decision-making tasks are associated with the measurement of attention and response inhibition and require selective attention to a particular stimulus with an appropriate response (e.g. Stroop test) (Al-Yahya et al, 2011). Lastly, working memory tasks are used to examine sustained attention and information processing speed, which require holding information in the mind that is available for processing for example the short-term memory recognition test (Al-Yahya et al, 2011).

There is growing interest in the interaction between cognitive and motor function in MS research that reflects its clinical importance, since 65% of individuals with MS report to have cognitive deficits (Chiaravalloti & DeLuca, 2008). The majority of the studies examining cognitive motor interference (CMI) (i.e. decline in performance in a motor task when a cognitive task is performed concurrently) have utilized walking speed and other spatiotemporal parameters as the main outcome to detect changes in motor performance (Hamilton et al, 2009; Wajda et al, 2013; Wajda & Sosnoff, 2015). Monticone et al. (2014) examined the reliability and measurement error of spatiotemporal gait parameters in pwMS during dual-task conditions. Gait was evaluated on two different occasions with a motor-cognitive task (walk while a word list generation (WLG) task was administered) and a motor-motor task (walk while carrying a tray with glasses) and it was reported that all gait parameters had good to excellent ICCs, but with the gait parameters of pwMS being slightly more variable between tests than healthy individuals (Monticone et al, 2014).

The majority of the literature focused on gait changes under dual-task conditions in pwMS with mild to higher EDSS scores. Only one study assessed gait performance in participants with CIS while performed a cognitive task (Kalron et al, 2010). It was been observed that gait parameters deteriorated even in this population when performing a cognitive task (i.e. modified WLG) compared to healthy individuals (Kalron et al, 2010). Several studies have utilized the WLG task while walking to examine CMI in pwMS (EDSS range: 2-6.5). Several authors reported that with the addition of this task, gait performance deteriorated (Sosnoff et al, 2011a; Wajda et al, 2013a; Wajda et al, 2013b; Motl et al, 2014). Other studies have found similar gait changes with the addition of a different cognitive task assessing mental tracking, such as alternating letters of the alphabet, serial 7 subtraction and counting backwards (Hamilton et al, 2009; Learmonth et al, 2014; Sandroff et al, 2015; Allali et al, 2016; Etemadi, 2016; Wajda et al, 2016).

Peruzzi et al. (2016) investigating the feasibility of a dual-task (DT) intervention program and pwMS underwent a six week virtual-reality (VR)-based treadmill training program. At baseline and at six weeks gait analysis was performed under single and dual-task conditions, with the serial 3 subtraction used as the cognitive task. It was observed that walking speed and stride length improved in the dual-task conditions after the intervention (Peruzzi et al, 2016).

A recent meta-analysis synthesized the evidence for differences in CMI between pwMS and healthy individuals. There was a small effect size indicating a non-significant difference in CMI between pwMS and healthy individuals, even though there was decrease in the walking performance in pwMS compared to the healthy individuals (Learmonth et al, 2017). The findings of this meta-analysis are consistent with the general

understanding that, regardless of the health status, undertaking a cognitive and a motor task simultaneously will result in decreased performance of the motor or cognitive task or a combination of both (Al-Yahya et al, 2011; Hamacher et al, 2015).

To our knowledge, although many studies have investigated dual tasking in pwMS, the influence of FES on the dual task cost has not been explored. In Chapter 6 the benefits of FES were explored under dual-task and fatiguing conditions in order to objectively document the effects of FES as perceived by pwMS.

2.6 Summary and rationale of the overall aims of the thesis

This background chapter has summarised and highlighted research findings and gaps in knowledge related to gait impairments and especially foot drop which is a key component of this thesis and reveals the following key points:

- A variety of outcome measures have been used to measure the effects of interventions to treat foot drop but there is been no systematic review evaluating both the quality of the methodology and quality of the evidence of studies assessing the psychometric properties of these outcomes.
- The psychometric properties of three-dimensional gait analysis, which is one of the most common outcomes to assess walking performance in MS, have not been established for the MS population.
- Currently, there is no objective evidence of the 'phenomenon/symptom' of fatigable foot drop in people with low EDSS levels (< 3.5), i.e. whose daily walking performance is not affected.

- The most common outcome measures used to assess the effects of interventions to treat foot drop, do not always reflect the benefits of these interventions noticed by pwMS which are linked to benefits in real life situations as opposed to walking tests in a lab. Thus, future studies should focus on the effect of interventions to treat foot drop in more real life conditions such as those requiring dual tasking.

Therefore, in accordance with the aforementioned bullet points, this PhD included the following research questions:

Study 1 (Chapter 3)

Research question 1: Which are the most frequently outcomes used to assess the effect of assistive technology to treat foot drop in pwMS?

Research question 2: What is the quality of published evidence on the psychometric properties of the outcome measures identified?

Research question 3: What is the level of evidence for the identified outcome measures?

Contribution: In collaboration with my supervisors, I was responsible for the conception and design of the study, performed the database searches, screening and analysis and preparing and submitting the manuscript reporting the results of this study for publication in a peer-reviewed journal.

Study 2 (Chapter 4)

Research question 1: What is the test-retest relative and absolute reliability of ankle kinematics and spatiotemporal parameters in two groups of pwMS with different level of walking impairments?

Research question 2: What is the intra-session relative and absolute reliability of ankle kinematics and spatiotemporal parameters in two groups of pwMS with different level of walking impairments?

Contribution: In collaboration with my supervisors, I have contributed to the conception and design of the study, data collection for the Group A, data analysis for both included groups and writing the study as an article for publication.

Study 3 (Chapter 5)

Research question 1: Can we objectively document gait deterioration, evidenced as foot drop, after a self-regulated exercise perturbation task in pwMS with minimal disability?

Research question 2: Are there any differences in gait characteristics between this group of pwMS and a healthy age-matched control group?

Contribution: Although this study was conceived an designed before prior to the start of my PhD, I was responsible for latter parts of ethical review (e.g. QMU ethics, amendments for NHS ethics), I was also responsible for all the data collection, data analysis and preparing the manuscript reporting the results of this study to be submitted in a peer-reviewed journal.

Study 4 (Chapter 6)

Research question 1: What are the differences in the direct orthotic effect of FES under dual-tasking and fatiguing conditions?

Research question 2: Is the dual-task cost of walking speed different between pwMS and an age-matched control group?

Research question 3: How is fatigability affecting the gait characteristics in pwMS and a healthy age-matched control group while simultaneously performing a cognitive task?

Contribution: In collaboration with my supervisors, I have contributed to the conception of the study, study design and acted as PI to gain NHS ethical approval, data collection and data analysis and preparing the manuscript reporting the results of this study to be submitted in a peer-reviewed journal.

Taking all these into consideration, the overall aim of this PhD thesis was to explore foot drop, its presence in pwMS with different disability levels and the psychometric properties of outcomes used to evaluate walking impairments.

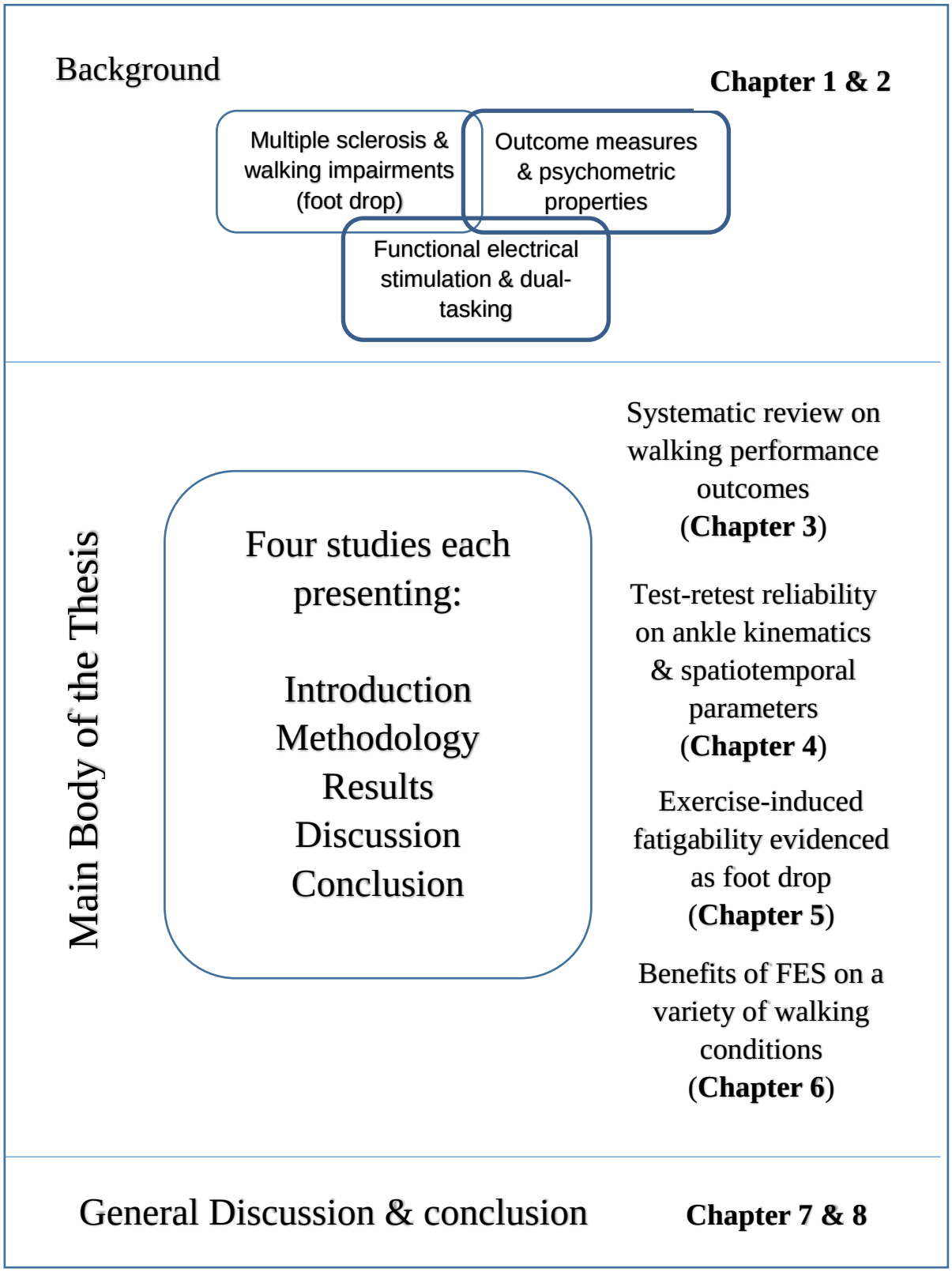


Figure 2.4 Overview of the thesis.

Chapter 3. Systematic review of the psychometric properties of walking performance measures

3.1 Purpose of the chapter

The purpose of this chapter is, through a systematic literature search, to identify all the outcome measures that have been used to assess the effects of assistive technology to treat foot drop. Further, to synthesize and evaluate the psychometric evidence of walking performance, effort of walking and lower limb function outcomes.

3.2 Introduction

One of the most common gait impairments is foot drop, which is the reduced dorsiflexion of the ankle during the swing phase of gait, potentially leading to trips or falls. Foot drop can be caused by weakness of the dorsiflexor muscles, impaired neural control causing co-contraction of agonist and antagonist muscles and increased tone in the plantarflexor muscles (Barret et al, 2009). In pwMS foot drop can also be caused by increased motor fatigability, which is described as the exercise-induced reduction in the ability of the muscles to produce force or power (Kluger et al, 2013). Two common interventions to treat foot drop are FES and AFOs. The most commonly used AFOs restrain the movement of the foot and thus reduce foot drop, but they do not allow active control of the ankle, which may result in an abnormal gait pattern (Bulley et al, 2015). On the contrary, FES involves electrical stimulation that is applied to the common peroneal nerve, eliciting the desired contraction to produce ankle dorsiflexion during the swing phase of gait. The

advantage of FES is that it facilitates a more normal gait pattern, increases walking speed and decreases the physiological cost of gait (Stein et al, 2006; Paul et al, 2008).

The effects of FES and AFOs on walking performance is currently evaluated via a wide variety of outcome measures including, for example, timed walking tests (e.g. 6MWT, 10mWT, T25FW) or patient or clinician reported instruments and rating scales [e.g. MSWS-12, HAI, Dynamic Gait Index (DGI)]. Instrumental motion analysis techniques are also used to objectively quantify the gait pattern. A comprehensive assessment of three-dimensional kinematics and kinetics can reveal minimal changes that cannot be observed visually (Bethoux & Bennett, 2011). For this reason, 3D gait analysis is widely used to discriminate between normal and abnormal gait patterns and to evaluate responses to interventions in a variety of populations, such as stroke (Stokic et al, 2009), cerebral palsy (Kainz et al, 2017a) and Parkinson's disease (Roiz et al, 2010; Pistacchi et al, 2017).

The outcome measures used to assess the efficacy of interventions such as assistive technology to treat foot drop need to be valid, reliable and responsive to change. Several studies have evaluated the psychometric properties of outcome measures used to assess the effects of ankle foot orthoses and functional electrical stimulation to treat foot drop [e.g. Goldman et al. (2008), Stellman et al. (2015), Learmonth et al. (2012, 2013)]. However, no systematic review exists that has evaluated both the evidence and the methodological quality of studies describing the psychometric properties of such outcome measures.

We, therefore, aimed to (i) identify studies that evaluated the effects of ankle foot orthoses and functional electrical stimulation in pwMS and then (ii) synthesize the available psychometric evidence for the designated subset of, walking performance, effort of walking and lower limb function, outcome measures identified. In so doing, we hoped to augment the evidence-base available to optimize the appropriate selection of outcome measure(s) to evaluate the efficacy of assistive technology to treat foot drop in pwMS.

3.3 Methods

3.3.1 First search: overview of outcome measures

The purpose of the first search of the literature was to identify those studies that assessed the effects of either FES or AFOs used to treat foot drop in pwMS. From these studies we identified the outcome measures used and the frequency of their use.

3.3.1.1 Search strategy and study selection

A comprehensive search of eight databases, including MEDLINE (1963-5/2017), CINAHL (1969-5/2017), EMBASE (1974-5/2017), SCOPUS (1963-5/2017), PsycINFO (1963-5/2017), AMED (1967-5/2017), SPORTDiscus (1963-5/2017) and Web of Science (1967-5/2017) was conducted in order to identify the articles that met the inclusion criteria. The search strategy included synonyms and keywords for functional electrical stimulation (e.g. 'Functional Electrical Stimulation', 'foot drop stimulation' and 'common peroneal stimulation') and ankle foot orthoses (e.g. 'Ankle Foot Orthoses' and 'splints')

and the population of interest (e.g. ‘multiple sclerosis’ and ‘demyelinating disease’). The full strategy can be found in Appendix 2.

The inclusion criteria for this search were: a) studies that have assessed the use of FES or AFOs to treat foot drop in pwMS and b) studies that included outcome measures that evaluate function, walking performance, fatigue and QoL. The exclusion criteria were: a) studies that used other forms of electrical stimulation (i.e. not functional) and those that evaluated orthoses for other joints than the ankle, b) studies that were reviews (i.e. systematic, meta-analysis, etc.), conference abstracts and editorials and c) studies in languages other than English, Greek or Dutch.

Two independent researchers (GA, MvdL) were involved in the screening of the articles for inclusion. After exclusion of irrelevant articles based on the titles and abstracts, the full-text of the remaining articles was examined for their eligibility. Reference lists of articles included in the review were searched for potentially relevant articles that were not retrieved in the original search. If any differences in opinion existed, consensus was made through discussion and a third reviewer (TM) was available if consensus between the primary two reviewers was not reached. From the eligible articles, we extracted the outcome measures that were employed to assess the effects of FES or AFOs and recorded the frequency of these measures being used.

3.3.2 Principal search: systematic review of the psychometric properties of outcome measures

The second and principal search was conducted to identify studies that evaluated the psychometric properties of outcome measures that assess walking performance, effort of walking and lower limb function in pwMS.

3.3.2.1 Search strategy and study selection

A similar protocol for the second search was followed as the one described above. A comprehensive search of MEDLINE (1976-5/2017), CINAHL (1995-5/2017), SCOPUS (1999-5/2017), EMBASE (1974-5/2017), PsycINFO (1963-5/2017), AMED (1967-5/2017), SPORTDiscus (1963-5/2017) and Web of Science (1967-5/2017) databases was conducted by combining the outcome measures of walking performance, effort of walking and lower limb function which were identified in the first search. The search strategy included keywords and synonyms of the population of interest (see first search), a subset of the identified outcome measures (e.g. '3D gait analysis', '10m walk test', etc.) and a search filter for identifying studies evaluating measurement properties, developed by Terwee et al. (2009a). The full search strategy is included as an appendix (Appendix 3).

The inclusion criteria for our second search were: studies that assessed the psychometric properties of a subset of the outcomes identified in the first search, namely those assessing walking performance, lower limb function and effort of walking. Although we acknowledge the importance of outcome measures such as QoL and fatigue, we decided to restrict the outcome measures in this review to those measures that are potentially

directly affected by the use of FES and AFOs. Further, the psychometric evidence for fatigue measures used in MS has been the subject of a previous review (Elbers et al, 2012). The exclusion criteria were: a) studies that were reviews (e.g. systematic and meta-analyses), abstracts from conferences or editorials, and b) full texts in peer reviewed journals published in languages other than English, Greek or Dutch. The procedures used to select the final set of papers were the same as those described for the first search.

3.3.3 Methodological quality

The methodological quality of the studies identified in the second search was assessed using the COSMIN. We chose the COSMIN checklist since is used to obtain a score for the methodological quality of a study evaluating one or more measurement properties of a particular outcome measure (Mokkink et al, 2012; Terwee et al, 2012). The COSMIN checklist has been assessed for the inter-rater agreement and reliability of each item, with the percentage agreement being appropriate, but the kappa coefficients for each item being relatively low (Mokkink et al, 2010a). However, to overcome low inter-rater agreement in scoring items, we familiarized with the grading process and developed specific guidelines as recommended by the developers of COSMIN. The COSMIN-checklist consists of nine boxes (internal consistency, reliability, measurement error, content validity, structural validity, construct validity, cross-cultural validity and responsiveness) with each box including 5-18 items. The reviewer selects the measurement properties evaluated in the study and scores the specific item-lists with ‘poor’, ‘fair’, ‘good’ or ‘excellent’ depending on the design and execution. The lowest score from the rated items determines the methodological quality of the measurement property (Mokkink et al,

2010b) (Appendix 4). Two reviewers (GA, MvdL) used the COSMIN checklist to rate the methodological quality of the measurement properties in all studies. Any disagreements in ratings were resolved through discussion.

As previously mentioned, in order to be consistent in our ratings we developed guidelines for the rating of specific questions/items in each of the measurement properties in the COSMIN checklist (Appendix 5). For example, all studies that used the EDSS as a comparator instrument were rated under the measurement property of construct validity, even if the authors stated that criterion validity was assessed. The questions for missing items and how they were handled was scored as 'not applicable' for measures that were not self-reported scales. For studies assessing within-day test-retest reliability, the items for patients being stable and the time interval being appropriate were rated as excellent.

The quality of the results of the psychometric properties of the outcome measures was assessed using the quality criteria by Terwee et al. (2007), which were recently revised by the authors (www.cosmin.nl, viewed online 2016). The quality of the results of the psychometric properties was rated as 'positive' (+), 'indeterminate' (?) or 'negative' (-) depending on the methods and results of the studies (Table 3.1).

Table 3.1 Quality criteria for measurement properties (Terwee et al, 2007; www.cosmin.nl, viewed online 2016).

Measurement property	**Rating***	**Criteria**
Reliability		
Internal consistency	+	At least limited evidence for unidimensionality or positive structural validity AND Cronbach's alpha(s) ≥ 0.70 and ≤ 0.95

	?	Not all information for '+' reported OR conflicting evidence for unidimensionality or structural validity OR evidence for lack of unidimensionality or negative structural validity
	-	Criteria for '+' not met
Reliability	+	ICC or weighted Kappa ≥ 0.70
	?	ICC or weighted Kappa not reported
	-	Criteria for '+' not met
Measurement error	+	SDC or LoA < MIC
	?	MIC not defined
	-	Criteria for '+' not met
Validity		
Construct validity (Hypothesis testing)	+	At least 75% of the results are in accordance with the hypotheses
	?	No correlations with instrument(s) measuring related construct(s) AND no differences between relevant groups reported
	-	Criteria for '+' not met
Criterion validity	+	Convincing arguments that gold standard is "gold" AND correlation with gold standard ≥0.70
	?	Not all information for '+' reported
	-	Criteria for '+' not met
Responsiveness		
Responsiveness	+	At least 75% of the results are in accordance with the hypotheses
	?	No correlations with changes in instrument(s) measuring related construct(s) AND no differences between changes in relevant groups reported
	-	Criteria for '+' not met

* + = positive rating; ? = indeterminate rating; - = negative rating

3.3.4 Data synthesis

The overall level of evidence for each outcome measure was reported according to the recommendations of the Cochrane Back Review Group. This overall score was given in relation to the methodological quality of the study and the results of the measurement properties. The evidence was rated as 'strong' (consistent (positive) findings in multiple studies of good methodological quality or in one study of excellent methodological quality), 'moderate' (consistent (positive) findings in multiple studies of fair methodological quality or in one study of good methodological quality), 'limited' (one study of fair methodological quality), 'conflicting' (both positive and negative findings), 'unknown' (only studies of poor methodological quality) (van Tulder et al, 2003). For instance, if the intra-rater reliability for a particular outcome measure had one study of poor quality and one of good quality showing positive results, the overall score was 'moderate'. Likewise, if there were four studies of fair methodological quality but only one with having a positive score for the quality of the results, the overall score was 'limited'.

3.4 Results

3.4.1 First search: overview of outcome measures

After a systematic search of the eight databases, we retrieved 1393 titles for screening according to our inclusion criteria (Figure 3.1). We retained 34 articles and identified 27 outcome measures evaluating lower limb function, walking performance, effort of

walking, fatigue and QoL. These outcomes measures were either self-reported measures (seven measures e.g. FSS, MSWS-12) or objective assessments (20 measures e.g. 6MWT, MSFC, spatiotemporal gait parameters). The most frequently used outcome measures were walking speed (mostly recorded over 10 meter distance), 3D gait kinematics and the Physiological Cost Index (PCI) (Figure 3.2).

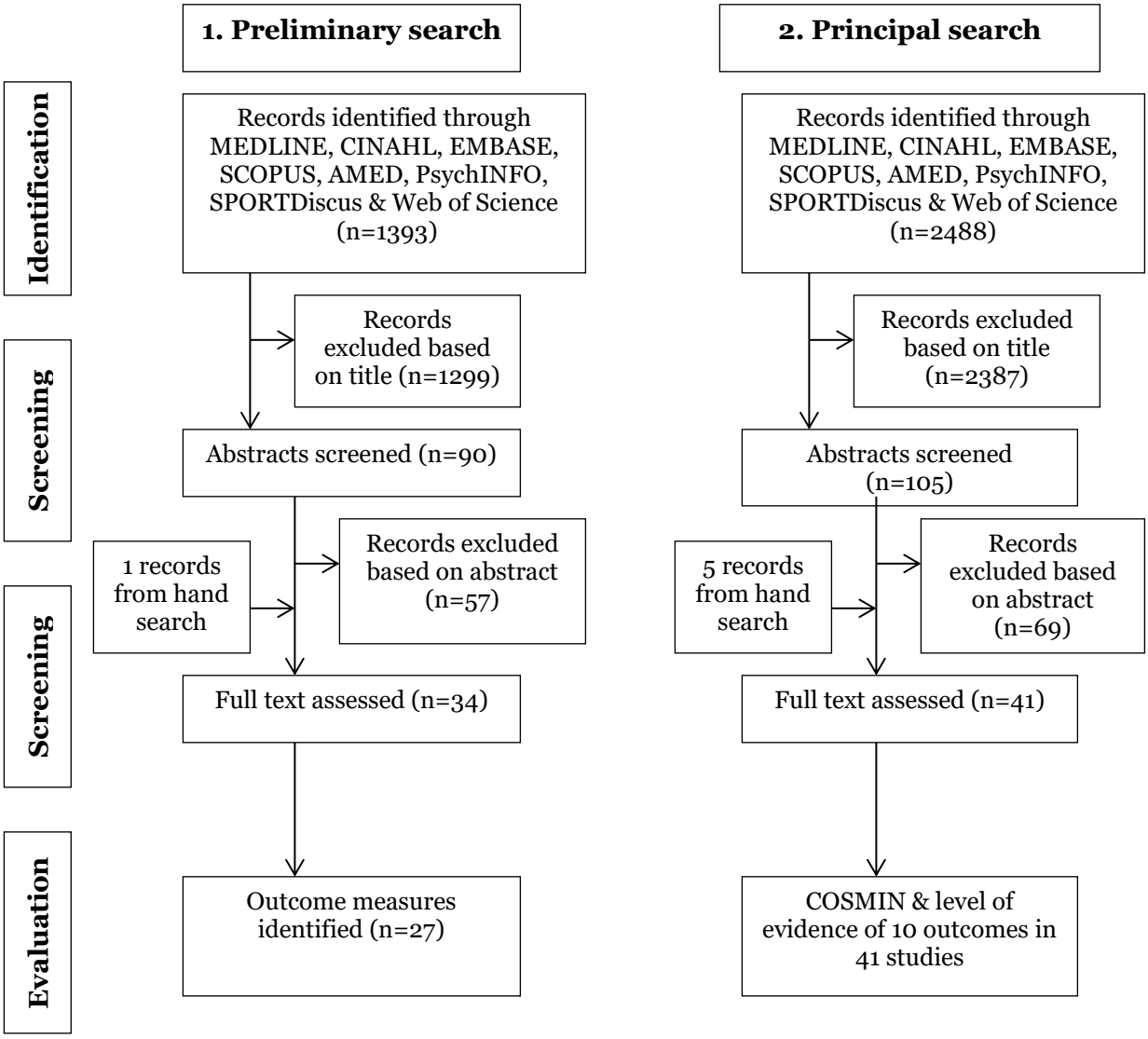


Figure 3.1 1. Preliminary search: identification of the outcomes measures that have been used to assess the effect of assistive technology for foot drop; 2. Principal search: studies evaluating the

psychometric properties of outcome measures of walking performance, effort of walking and lower limb function.

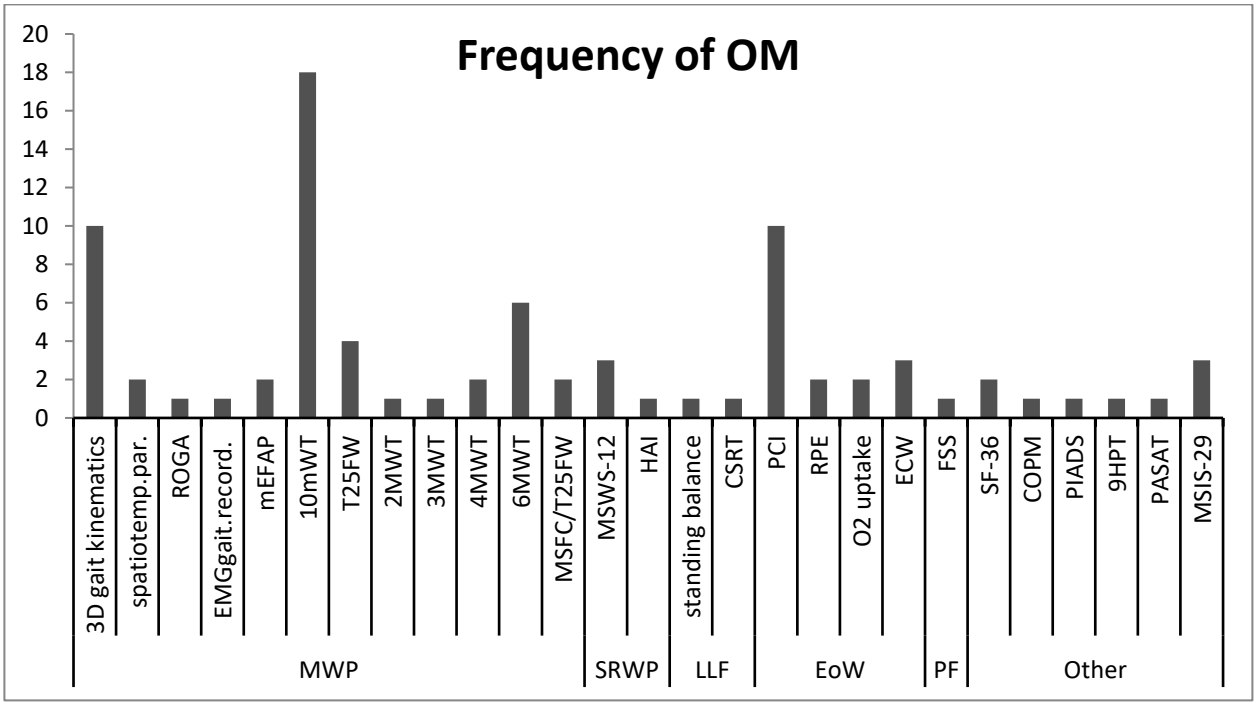


Figure 3.2 Outcome measures identified in the preliminary search and the reported frequency of use.

Abbreviations: MWP: measured walking performance; SRWP: self-reported walking performance; LLF: lower limb function; EoW: effort of walking; PF: perceived fatigue; spatiotemp. par.: spatiotemporal parameters; ROGA: Rivermead Observational Gait Analysis; EMG gait record: electromyography gait recording; mEFAP: modified Emory Functional Ambulation Profile; 10mWT: 10 meter Walk Test; T25FW: Timed 25 Foot Walk; 2MWT: 2 Minute Walk Test; 3MWT: 3 Minute Walk Test; 4MWT: 4 Minute Walk Test; 6MWT: 6 Minute Walk Test; MSFC: Multiple Sclerosis Functional Composite; MSWS-12: Multiple Sclerosis Walking Scale-12; HAI: Hauser Ambulation Index; CSRT: choice stepping reaction time; PCI: Physiological Cost Index; RPE: Rate of Perceived Exertion; ECW: energy cost of walking; FSS: Fatigue Severity Scale; SF-36: 36-Item Short Form Health Survey; COPM: Canadian Occupational Performance Measure; PIADS: Psychosocial Impact of Assistive Devices Scale; 9HPT: 9-Hole Peg Test; PASAT: Paced Serial Addition Test; MSIS-29: Multiple Sclerosis Impact Scale.

3.4.2 Principal search: systematic review of the psychometric properties of outcome measures

3.4.2.1 Description of included studies

The systematic search of eight databases resulted in the identification of 2488 potentially relevant titles. After independent screening according to the inclusion and exclusion criteria, we retained 36 articles with further reference citation tracking resulting in five

additional articles (Figure 3.1). Four studies (Hoogervost et al, 2002; Kragt et al, 2006; Goldman et al, 2013; Motl et al, 2014) were excluded at the full-text screening stage because they aimed at validating a previously reported MCID or cut-off points for a certain outcome measure and did not validate the outcome measure itself. Although of interest, the methodology of these studies is different from those reporting the psychometric properties of the outcome measures themselves and are therefore not appropriate to be assessed using the COSMIN checklist and Terwee criteria. In total, we included 41 articles reporting the psychometric properties of 10 outcome measures [MSFC, MSWS-12, spatiotemporal parameters, 10mWT, T25FW, 2MWT, 6MWT, RPE, peak oxygen uptake (VO_2 peak) & reaction time/movement time (RT/MT)] which all have been used to assess the effects of assistive technology to treat foot drop. Using the COSMIN taxonomy the following measurement properties were evaluated: reliability was assessed in 18 studies [(intra-rater n=3; inter-rater n=3; test-retest n=14), 8 outcome measures], measurement error in four studies (six outcome measures) and internal consistency in six studies (one outcome measure). Hypothesis testing/construct validity was evaluated in 15 studies (nine outcome measures) and responsiveness was assessed in 15 studies (seven outcome measures). Most studies assessed the MSWS-12 (n=12), followed by the 6MWT (n=11) and the T25FW (n=11). The agreement between the two raters (GA & MvdL) in the items of all the measurement properties was 94.8% and for the final scores of each property the agreement was 94.7%. Upon discussion, any disagreement regarding the rating of the items or the total score of each property was resolved. Studies included pwMS with RR, SP, PP and CIS with EDSS levels ranging from 0-8.5, with some studies not reporting this information (Kaufman et al, 2000; Hobart et al, 2003; Motl et al, 2011; Freeman et al,

2013; Toomey et al, 2013). The majority of the studies reported a mean of EDSS of four or more and five studies reported a mean EDSS of six (Vaney et al, 1996; Sosnoff et al, 2011b; Coleman et al, 2012; Learmonth et al, 2012; Hobart et al, 2013). The sample size was 6796 in total for the 41 studies, with the number of females (n=2109) exceeding the number of males (n=972) and with some studies not reporting the gender of the participants (Cutter et al, 1999; Kaufman et al, 2000; Kragt et al, 2008; Motl et al, 2011). Table 3.2 presents an overview of the results together with the COSMIN rating and the rating of the quality of the results according to the revised Terwee criteria (Terwee et al, 2007; Mokkink et al, 2010b; www.cosmin.nl, viewed online 2016).

3.4.3 Methodological quality and strength of evidence

3.4.3.1 Reliability

The methodological quality of the studies was rated according to the COSMIN checklist as 'good' (n=3) (Cohen et al, 2001; Hobart et al, 2003; Learmonth et al, 2013a), as 'fair' (n=3) (Goldman et al, 2008; Motl et al, 2011; Larson et al, 2013) and 'poor' (n=12) (Vaney et al, 1996; Cohen et al, 2000; Kaufman et al, 2000; Schwid et al, 2002; Paltamaa et al, 2005; Fry et al, 2006; Learmonth et al, 2012; Hobart et al, 2013; Toomey et al, 2013; Feys et al, 2014; Heine et al, 2015; Cleland et al, 2016). The main reasons for a lower score included not reporting the ICC or weighted Kappa, not describing the ICC model used, small sample size and the lack of an explicit statement that the repeated measurements were independent. Using the revised Terwee quality criteria (Terwee et al, 2007; www.cosmin.nl, viewed online 2016), the evidence for reliability in 13 studies (seven

outcome measures) (Cohen et al, 2000; Cohen et al, 2001; Hobart et al, 2003; Paltamaa et al, 2005; Fry et al, 2006; Goldman et al, 2008; Motl et al, 2011; Learmonth et al, 2012, Learmonth et al, 2013a; Larson et al, 2013; Toomey et al, 2013; Heine et al, 2015; Cleland et al, 2016) were rated as 'positive' and the remaining five (four outcome measures) (Vaney et al, 1996; Kaufman et al, 2000; Schwid et al, 2002; Hobart et al, 2013; Feys et al, 2014) were rated as 'indeterminate' because neither ICC nor weighted Kappa were reported. From the eight outcome measures that were evaluated for reliability (intra- & inter-rater, test-retest), seven of them demonstrated good and excellent values of ICC ranging from 0.86-0.96 and only for RPE the ICC values were moderate (0.706).

3.4.3.2 Measurement error

Of the four studies that evaluated measurement error (six outcome measures), the methodological quality of three (Paltamaa et al, 2005; Learmonth et al, 2012; Heine et al, 2015) was rated as 'poor' due to a small sample size ($n < 30$) and due to testing conditions not being similar. The methodology in one study (Learmonth et al, 2013a) was rated as 'fair' because it was unclear whether the patients were stable in the interim period. The quality of the results for measurement error in all four studies was rated as indeterminate ('?') because in none of the studies the Minimal Important Change (MIC) values was reported, which is required to interpret whether the measurement error is acceptable (de Vet et al, 2006a).

3.4.3.3 Internal consistency

There were six studies that evaluated internal consistency. The methodological quality of four (Hobart et al, 2003; Motl et al, 2011; Engelhard et al, 2016; Mokkink et al, 2016) was rated as ‘excellent’, for one (Motl et al, 2008) it was rated as ‘good’ and one (Motl et al, 2010) as ‘poor’ due to a small sample size. All six studies evaluated the MSWS-12 and were rated as positive (‘+’) for the quality of their results.

Table 3.2 Summary of the study characteristics, rating of the methodological quality using the COSMIN guidelines and rating of the quality of the results using the Terwee criteria (Terwee et al, 2007; Mokkink et al, 2010b; www.cosmin.nl, viewed online 2016).

Author/Year	Patient characteristics	COSMIN Measurement Property	Results	Rating	
				Methodological quality	Quality of the results
Choice reaction time & movement time (RT/MT)					
Apache et al (2002)	n = 178 ,RR, SP,PP EDSS 0-6.5	Hypothesis testing	r_s = 0.84 with EDSS	**Fair** [a]	+
Apache et al (2005)	n = 40, RR, SP EDSS median 4.5 3 sessions in 1-year	Responsiveness (no intervention)	RT/MT mean change =16.6% (.1)	**Poor** [b]	**?**
MSFC					
Cohen et al (2000)	n = 10, SP EDSS mean 5.2 6 sessions (2 per day over 2 weeks)	Reliability	Intra-rater: ICC = 0.97 (session 4-5) Inter-rater: ICC = 0.96 (session 7-8)	**Poor** [c]	+
Cohen et al (2001)	n = 436, SP EDSS mean 5.2	Reliability Hypothesis testing	Intra-rater: ICC (over 4 sessions) = 0.87 r_s = -0.56 with EDSS	**Good** [d, e] **Fair** [o, q]	+ +

	3 pre-baseline sessions over 28 days			**Methodological quality**	**Quality of the results**
Cutter et al (1999)	n = 378, RR, SP EDSS 0-6.5 3 annual sessions	Hypothesis testing Responsiveness (no intervention)	r_s = -0.22 with EDSS Average composite change Z-score: Baseline = -0.07 1-year = -0.07 2-year = -0.16	**Fair** [a, o ,q] **Fair** [q]	**?** **?**
Hobart et al (2004)	n = 133, RR, SP,PP EDSS mean 3.1	Hypothesis testing	r = -0.64 with EDSS	**Good** [f, w]	+
Kalkers et al (2001)	n = 131, RR, SP,PP EDSS mean 3.1	Hypothesis testing	r_s = -0.25 with T2 lesion load r_s = -0.24 with T1 lesion load	**Poor** [h]	**?**
Kragt et al (2008)	n = 161, PP EDSS mean 5.0	Responsiveness (no intervention)	ES: EDSS = 0.23 MSFC = 0.16	**Poor** [b]	**?**
Miller et al (2000)	n = 300 EDSS 0-8.5	Hypothesis testing	r_s = -0.80 with HRQoL	**Poor** [h]	**?**
MSWS-12					
Baert et al (2014)	n = 284, RR, SP,PP EDDS mean 4.8	Responsiveness	AUC with Global Rating Scale: Whole group = 0.73	**Fair** [a]	+

	2 sessions (pre & post)	(physical rehabilitation)	EDSS≤4 = 0.64 EDSS 4.5-6.5 = 0.77		
Filipović Grćić et al (2011)	n = 49, RR EDSS mean 3.0 2 sessions (pre & post)	Responsiveness (IVMP for 1month)	SRM = 1.05 ES = 1.02 RE (%) = 82.4	**Fair** [a]	**?**
Freeman et al (2013)	n = 70, RR, SP, PP 3 annual sessions	Responsiveness (no intervention)	ES = -0.07 SEM = 5.66 r <0.35 with walking speed & RMI	**Fair** [j]	**?**
Hobart et al (2003)	Community sample: n = 602 2 sessions (10 days apart)	Internal consistency	Community sample: Cronbach's α = 0.97 PPMS sample: Cronbach's α = 0.97 Steroids sample: Cronbach's α = 0.94	**Excellent**	**+**
		Reliability	Community sample: Test-retest ICC = 0.94	**Good** [d, e, k]	**+**
	Hospital-based sample: PP MS = 78 Steroids = 54 2 sessions (6 weeks apart)	Hypothesis testing	Steroids sample: r_s = 0.65 with EDSS	**Fair** [i, q]	**+**
		Responsiveness (steroid treatment)	With EDSS: ES = 0.45 SRM = 0.45 RE = 0.31	**Fair** [m, q, t]	**?**
	n = 82, RR, SP, PP	Reliability	Test-retest: ICC(2,1) = 0.93	**Good** [k, t]	**+**

Learmonth et al (2013)	EDSS mean 3.5 2 sessions (7 days apart)	Measurement error	SEM = 8; CV (5) = 27 MDC_{95} = 22; $\%MDC_{95}$ = 53%	**Fair** [i, m, n]	**?**
McGuigan et al (2004)	Community sample = 149	Hypothesis testing	r_s = 0.84 with EDSS	**Fair** [a, o, q]	**?**
	Outpatient sample = 53 RR, SP, PP EDSS mean 4.0 2 sessions	Responsiveness (no intervention)	Z-score = -2.87	**Poor** [b]	**?**
Motl et al (2008)	n = 133, RR, SP, PP EDSS mean 4.9 1 session	Internal consistency	Cronbach's α =.97	**Good** [p, x]	**+**
		Hypothesis testing	r_s = .77 with MSIS-29 (physical) r_s = .36 with MSIS-29 (psychological) r_s = .80 with EDSS	**Fair** [q]	**+**
Motl et al (2010)	n = 24, RR PDDS median 1.0 1 session	Internal consistency	Cronbach's α = .95	**Poor** [c]	**+**
		Hypothesis testing	With O_2 cost of walking at: CWS, r = 0.64 FWS, r = 0.61	**Poor** [c]	**+**

			SWS, r = 0.64 With O_2 consumption: CWS, r = 0.24 FWS, r = 0.14 SWS, r = 0.44		
Motl et al (2011)	n = 269, RR 3 sessions over a year	Internal consistency	Cronbach's α: Baseline = .96 6-month = .97 12-month = .97	**Excellent**	+
		Reliability	Test-retest ICC(2): Across 6-months = .86 Across 12-months = .87	**Fair** [i]	+
Pilutti et al (2013)	n = 268, RR, SP,PP PDDS median 3.0 1 session	Hypothesis testing	r_s = .72 with T25FW r_s = -.75 with 6MWT	**Fair** [q]	+
Mokkink et al (2016)	n = 625, RR, SP, PP,PR, CIS EDSS median 3.5	Internal consistency	RMSEA = 0.078 CFI = 1.000 TLI = 0.999 SRMR = 0.019 Guttman's lambda2 = 0.98	**Excellent**	+

Engelhard et al (2016)	n = 293, RR, SP, PP,PR	Internal consistency	1D Rasch: BIC = 6112.5; AIC = 5933.7 3D GRM: BIC = 5972.7; AIC = 5677.3	**Excellent**	+
FAP/ Spatiotemporal parameters					
Sosnoff et al (2011)	n = 13, RR, SP EDSS median 6.0 1 session	Hypothesis testing	FAP: r_s = -0.82 with T25FW r_s = -0.49 with MSWS-12 r_s = -0.81 with EDSS	**Poor** [c]	+
Pilutti et al (2013)	n = 268, RR, SP,PP PDDS median 3.0 1 session	Hypothesis testing	Speed with T25FW: r = -.68 Cadence with T25FW: r = -.50 Speed with 6MWT: r = .67 Cadence with 6MWT: r = .52	**Fair** [q]	+
10mWT					
Feys et al (2014)	n = 102, RR, SP,PP EDSS mean 4.6 3 sessions within a day	Reliability	Test-retest: Within-day variability (%) at usual speed: -Community walkers (CW) = 22.6 -Limited CW = 26.6 -Most limited CW = 43.3	**Poor** [b]	**?**

			Within-day variability (%) at fastest speed: -CW = 12.1 -Limited CW = 23.4 -Most limited CW = 38.4		
Freeman et al (2013)	n = 70, RR, SP, PP 3 annual sessions	Responsiveness (no intervention)	ES = 0.001 r <0.35 with MSWS-12 & RMI	**Fair** [j]	**?**
Kempen et al (2011)	n = 156 , RR EDSS mean 2.5 6 sessions in 6 years	Responsiveness (no intervention)	AUC = 0.79 with MFWC6 AUC = 0.86 with MFWC5 AUC = 0.74 with MFWC4 AUC = 0.82 with MFWC3	**Fair** [j]	**+**
Paltamaa et al (2005)	Test-retest n = 19 Inter-rater n = 9 RR, SP, PP EDSS 0-6.5 2 sessions (1 week apart)	Reliability Measurement error	Test-retest: ICC = 0.91 Inter-rater: ICC = 0.93 Test-retest: SEM = 0.09m/s Inter-rater: SEM = 0.10m/s	**Poor** [c] **Poor** [c]	**+** **?**
Paltamaa et al (2008)	Baseline n = 120 Follow-up n = 109 RR, PP EDSS median 2.0	Responsiveness (no intervention)	AUC = 0.76 with EDSS $MIC_{deterioration}$ = -0.19	**Fair** [a]	**+**

	3 sessions in 2years				
Stellman et al (2015)	n = 28, RR, SP, PP EDSS mean 3.2 1 session	Hypothesis testing	r = 0.61 with accelerometry	**Poor** [c]	+
Vaney et al (1996)	Reliability n = 25 Responsiveness n = 115 EDSS mean 6.6 5 sessions within-day	Reliability Responsiveness (physical & occupational therapy)	Test-retest: r_s = -0.8 with RMI Not adequate statistical information for responsiveness	**Poor** [c, u] **Poor** [y]	**?** **?**
Timed 25-Foot Walk					
Baert et al (2014)	n = 284, RR, SP,PP EDDS mean 4.8 2 sessions (pre & post)	Responsiveness (physical rehabilitation)	AUC with Global Rating Scale: Whole group = 0.50 EDSS≤4 = 0.64 EDSS 4.5-6.5 = 0.45	**Fair** [a]	+
Coleman et al (2012)	n = 296, RR, RP, SP, PP EDSS mean 5.8 4 sessions	Responsiveness (dalfampridine treatment)	r_s = -0.39 with CGI MICD = 0.35 m/s Relative improvement = 17.2%	**Fair** [a]	+

Filipović Grćić et al (2011)	n = 49, RR EDSS mean 3.0 2 sessions (pre & post)	Responsiveness (IVMP for 1 month)	SRM = 0.55 ES = 0.27 RE (%) = 68.3	**Fair** [a, t]	**?**
Hobart et al (2013)	n = 533, RR, SP,PP EDSS mean 6.0 9 sessions	Reliability Hypothesis testing	Variability ranged from 10.03 – 11.44 r = -0.20 to -0.43 with MSWS-12	**Poor** [b] **Excellent**	**?** **+**
Kaufman et al (2000)	n = 133, SP 3 sessions (6 month period)	Reliability Responsiveness (no intervention)	Not adequate statistical information for reliability Not adequate statistical information for responsiveness	**Poor** [b] **Poor** [b, h]	**?** **?**
Larson et al (2013)	n = 36, RR EDSS mean 3.5 2 sessions 1 week apart	Reliability	Test-retest ICC = 0.92	**Fair** [c]	**+**
Learmonth et al (2012)	n = 24 EDSS mean 6.02 2 sessions 1 week apart n = 82, RR, SP, PP	Reliability Measurement error Reliability	Test-retest ICC(2,3) = 0.94 SEM = 4.56s MDC_{95} = 12.6s Test-retest ICC(2,1) = 0.991	**Poor** [c] **Poor** [c] **Good** [t]	**+** **?** **+**

Learmonth et al (2013)	EDSS mean 3.5 2 sessions (7 days apart)	Measurement error	SEM = 1s MDC_{95} = 2.7s % MDC_{95} = 36	**Fair** [i, m, n]	**?**
Schwid et al (2002)	n = 63 EDSS 0-6.5 5 sessions	Reliability	Test-retest reliability: 95% CI: ± 16% of patients baseline score	**Poor** [b]	**?**
van Winsen et al (2010)	n = 112, CIS, RR, SP, PP EDSS mean 4.5 2 sessions (pre & post)	Responsiveness (IVMP for 6 weeks)	Sensitivity (%) = 25 Specificity (%) = 90 LR+ = 2.50 LR- = 0.83	**Fair** [a]	**?**
Jensen et al (2016)	n = 105 EDSS mean 5.6 2 sessions	Responsiveness (SR-Fampridine treatment)	MCID = 1.3s %MCID = 14.2	**Poor** [h, y]	**?**
2-Minute Walk Test					
Baert et al (2014)	n = 284, RR, SP, PP EDDS mean 4.8 2 sessions (pre & post)	Responsiveness (physical rehabilitation)	AUC with Global Rating Scale: Whole group = 0.64 EDSS≤4 = 0.74 EDSS 4.5-6.5 = 0.60	**Fair** [a]	**+**

Feys et al (2014)	n = 102, RR, SP,PP EDSS mean 4.6 3 sessions within a day	Reliability	Within-day variability (%): CW = 12.0 Limited CW = 13.8 Most limited CW = 26.3	**Poor** [b]	**?**
Filipović Grćić et al (2011)	n = 49, RR EDSS mean 3.0 2 sessions (pre & post)	Responsiveness (IVMP for 1 month)	SRM = 0.89 ES = 0.54 RE (%) = 95.1	**Fair** [a]	**?**
Stellman et al (2015)	n = 28, RR, SP, PP EDSS mean 3.2 1 session	Hypothesis testing	r = 0.79 with accelerometry	**Poor** [c]	**+**
6-Minute Walk Test					
Baert et al (2014)	n = 284, RR, SP,PP EDDS mean 4.8 2 sessions (pre & post)	Responsiveness (physical rehabilitation)	AUC with Global Rating Scale: Whole group = 0.68 EDSS≤4 = 0.77 EDSS 4.5-6.5 = 0.65	**Fair** [a]	**+**
Feys et al (2014)	n = 102, RR, SP,PP EDSS mean 4.6 3 sessions within a day	Reliability	Within-day variability (%): CW = 10.1 Limited CW = 15.7 Most limited CW = 28.7	**Poor** [b]	**?**

Freeman et al (2013)	n = 70, RR, SP, PP 3 annual sessions	Responsiveness (no intervention)	ES = 0.03 'general mobility': r = 0.499	**Fair** [j]	**?**
Fry et al (2006)	n = 12, RR, SP, PP EDSS mean 3.6 2 sessions (1 week apart)	Reliability	Test-retest ICC = 0.96	**Poor** [c]	**+**
Goldman et al (2008)	n = 40, RR, SP, PP EDSS 0-6.5 3 sessions (in 4 hours)	Reliability	Test-retest: ICC = 0.94 Inter-rater: ICC = 0.91	**Fair** [t]	**+**
Learmonth et al (2012)	n = 24 EDSS mean 6.02 2 sessions 1 week apart	Reliability Measurement error	Test-retest: ICC (2,1) = 0.96 SEM = 27.48m MDC_{95} = 76.2m	**Poor** [c] **Poor** [c]	**+** **?**
Learmonth et al (2013)	n = 82, RR, SP, PP EDSS mean 3.5 2 sessions (7 days apart)	Reliability Measurement error	Test-retest: ICC(2,1) = 0.96 SEM = 32m MDC_{95} = 88m % MDC_{95} = 20	**Good** [t] **Fair** [i, m, n]	**+** **?**
Paltamaa et al (2005)	Test-retest n = 19 Inter-rater n = 9	Reliability	Test-retest: ICC = 0.96 Inter-rater: ICC = 0.93	**Poor** [c]	**+**

	RR, SP, PP EDSS 0-6.5 2 sessions (1 week apart)	Measurement error	Test-retest: SEM = 30.65 m Inter-rater: SEM = 35.85 m	**Poor** [c]	**?**
Paltamaa et al (2008)	Baseline n = 120 Follow-up n = 109 RR, PP EDSS median 2.0 3 sessions in 2 years	Responsiveness (no intervention)	AUC = 0.76 with EDSS $MIC_{deterioration}$ = -55.06	**Fair** [a]	**+**
Stellman et al (2015)	n = 28, RR, SP, PP EDSS mean 3.2 1 session	Hypothesis testing	r = 0.68 with accelerometry	**Poor** [c]	**+**
Toomey et al (2013)	n = 8 1 session(4assessors)	Reliability	Inter-rater: ICC = 0.984	**Poor** [c]	**+**
RPE					
Heine et al (2015)	n = 31 RR, SP, PP EDSS mean 2.5 2 sessions (1-3 weeks apart)	Reliability Measurement error	Test-retest: ICC = 0.706 SEM = 1.1 $SDC_{individual}$ = 2.9 SDC_{group} = 0.52 LoA = -2.9-2.9	**Poor** [z] **Poor** [z]	**+** **?**

Cleland et al (2016)	n = 16 RR, SP, PP EDSS median 1.75 2 sessions (6-10 days apart)	Reliability Hypothesis testing	Test-retest: ICC = 0.870 r = .691 with VO_2 (L/min) r = .507 with VO_2 (mL/kg/min)	**Poor** [c] **Poor** [c, h]	+ **?**
VO_2 peak					
Heine et al (2015)	n = 31 RR, SP, PP EDSS mean 2.5 2 sessions (1-3 weeks apart)	Reliability Measurement error	Test-retest: ICC = 0.933 for VO_2 peak ($mL \cdot kg^{-1} \cdot min^{-1}$) VO_2 peak ($mL \cdot kg^{-1} \cdot min^{-1}$): SEM = 1.7 $SDC_{individual}$ = 4.6 SDC_{group} = 0.82 LoA = -5.0-4.3	**Poor** [z] **Poor** [z]	+ **?**

COSMIN item rating: a: hypothesis vague or not formulated, possible to deduce; b: not appropriate statistical methods; c: small sample size; d: no description of ICC model used; e: assume that patients were stable in the interim period; f: expected magnitude of the correlations not stated; g: assumable that statistical methods were appropriate; h: no information about the psychometric properties of the comparator instruments; i: not clear how missing items were handled; j: unclear or not described what occurred in the interim period; k: assumable that measurements were independent; l: AUC or correlations not calculated; m: unclear if patients were stable; n: doubtful whether time interval was appropriate; o: poor description of the comparator instrument; p: no description of the % of missing data; q: some information on measurement properties or a reference; r: internal consistency not calculated for each subscale separately; s: no ICC, Spearman or Pearson's correlations calculated; t: due to sample size; u: only correlations, not ICC

calculated; v: minimal number of hypothesis formulated a priori; w: expected direction of the correlations or differences not stated; x: not described but can be deduced how missing items were handled; y: unclear what was expected; z: test conditions were not similar

Abbreviations: AIC: Akaike information criterion; AUC: Area under the curve; BIC: Bayesian information criterion; CFI: comparative fit index; CGI: Clinician Global Impression; CIS: Clinically Isolated Syndrome; CWS: comfortable walking speed; ES: effect size; FWS: faster walking speed; GPCM: generalized partial credit model; GRM: graded response model; HRQoL: Health-related Quality of Life; ICC: Intraclass Correlation Coefficient; IVMP: intravenous methylprednisolone therapy; LoA: limits of agreement; LR: Likelihood ratio; MDC: minimum detectable change; MIC: minimal important change; MICD: Minimally important clinical difference; MFWC: Modified Functional Walking Categories; PDDS: Patient Determined Disease Steps; PP: Primary Progressive; r: Pearson's correlations; r_s: Spearman coefficient; RE: relative efficiency; RMI: Rivermead Mobility Index; RMSEA: root mean square error of approximation; RR: Relapsing Remitting; SDC: smallest detectable change; SEM: standard error of mean; SP: Secondary Progressive; SRM: standardized response mean; SRMR: root mean square residual; SWS: slower walking speed; TLI: Tucker-Lewis index.

3.4.3.4 Hypothesis testing/construct validity

Fifteen studies assessed construct validity, with only one (Hobart et al, 2013) with an 'excellent' methodological quality and one in which was rated as 'good' (Hobart et al, 2004). The methodological quality of seven (Cutter et al, 1999; Cohen et al, 2001; Apache et al, 2002; Hobart et al, 2003; McGuigan et al, 2004; Motl et al, 2008; Pilutti et al, 2013) was rated as 'fair' either because the hypotheses were vague or due to limited information regarding the comparator instruments and its psychometric properties. The other studies (Miller et al, 2000; Kalkers et al, 2001; Motl et al, 2010; Sosnoff et al, 2011b; Stellman et al, 2015; Cleland et al, 2016) were rated as 'poor' due to a small sample size or the absence of information regarding the comparator instruments. Applying the Terwee quality criteria (Terwee et al, 2007; www.cosmin.nl, viewed online 2016), the quality of the results reported in 10 studies (Cohen et al, 2001; Apache et al, 2002; Hobart et al, 2003; Hobart et al, 2004; Motl et al, 2008; Motl et al, 2010; Sosnoff et al, 2011b; Hobart et al, 2013; Pilutti et al, 2013; Stellman et al, 2015) was rated as positive ('+') and in five studies (Cutter et al, 1999; Miller et al, 2000; Kalkers et al, 2001; McGuigan et al, 2004; Cleland et al, 2016) as indeterminate ('?') as the correlations presented were with unrelated constructs. The construct validity of seven of the studies reporting on six laboratory based measures and one self-perceived scale of walking performance used the EDSS as a comparator instrument. The comparator instrument in other studies were outcome measures such as the MSWS-12, Multiple Sclerosis Impact Scale-29 (MSIS-29), accelerometry and O_2 cost of walking. Table 3.2 includes information regarding the comparator instruments and correlation coefficients presented in studies assessing construct validity.

3.4.3.5 Responsiveness

Responsiveness was evaluated in 15 studies, with the methodological quality of nine (Cutter et al, 1999; Hobart et al, 2003; Paltamaa et al, 2008; van Winsen et al, 2010; Filipović Grćić et al, 2011; Kempen et al, 2011; Coleman et al, 2012; Freeman et al, 2013; Baert et al, 2014) rated as 'fair' and the remainder classed as 'poor' (Vaney et al, 1996; Kaufman et al, 2000; McGuigan et al, 2004; Apache et al, 2005; Kragt et al, 2008; Jensen et al, 2016). Most of the studies had a vague hypothesis or did not use appropriate statistical methods and this lowered their rating. Only four studies (Paltamaa et al, 2008; Kempen et al, 2011; Coleman et al, 2012; Baert et al, 2014) investigating the MSWS-12, 10mWT, T25FW, 2MWT and 6MWT received a 'positive' rating and the remaining 11 (Vaney et al, 1996; Cutter et al, 1999; Kaufman et al, 2000; Hobart et al, 2003; McGuigan et al, 2004; Apache et al, 2005; Kragt et al, 2008; van Winsen et al, 2010; Filipović Grćić et al, 2011; Freeman et al, 2013; Jensen et al, 2016) were rated as 'indeterminate' due to correlations with unrelated constructs or the lack of differences between relevant groups. Of the 15 studies evaluating responsiveness, only two studies (Paltamaa et al, 2008; Coleman et al, 2012) reported on the MCID (MIC) for the 10mWT, T25FW and 6MWT.

3.4.4 Level of evidence – data synthesis

The overall levels of evidence for the psychometric properties of each outcome measure are summarized in Table 3.3. It was found that the MSWS-12 has strong positive evidence for its internal consistency and test-retest reliability, moderate positive evidence for its

construct validity when compared to MSIS-29 and O_2 cost of walking and limited positive evidence for its responsiveness. The MSFC showed moderate positive evidence for its intra-rater reliability and construct validity, while for the remaining measurement properties, including responsiveness, the evidence was 'unknown'. For lower limb reaction/movement time, there was limited positive evidence for construct validity, but for responsiveness the evidence was 'unknown'. Strong evidence was found for the construct validity of the T25FW while for responsiveness and for test-retest reliability the evidence was moderately positive. Spatiotemporal parameters were classed as having a limited positive level of evidence for construct validity. For the 10mWT the level of evidence for its responsiveness was moderately positive, while for the other measurement properties this was 'unknown'. Limited positive evidence was found for the responsiveness of the 2MWT. For the 6MWT, the level of evidence for responsiveness and test-retest reliability was moderately positive, while the evidence for the inter-rater reliability was limited positive. The level of evidence for the measurement properties assessed for VO_2 peak and RPE were all 'unknown'.

Table 3.3 Level of evidence for each outcome measure identified in the principal search.

Outcome measure	Internal consistency	Reliability **Intra-rater**	**Inter-rater**	**Test-retest**	Measurement error	Hypothesis testing	Responsiveness
RT/MT	n/a	n/a	n/a	n/a	n/a	+ Limited	? Unknown

MSWS-12	+++ Strong	n/a	n/a	+++ Strong	? Unknown	++ Moderate	+ Limited
MSFC	n/a	++ Moderate	? Unknown	n/a	n/a	++ Moderate	? Unknown
Spatiotemporal parameters	n/a	n/a	n/a	n/a	n/a	+ Limited	n/a
10mWT	n/a	n/a	? Unknown	? Unknown	? Unknown	? Unknown	++ Moderate
T25FW	n/a	n/a	n/a	++ Moderate	? Unknown	+++ Strong	++ Moderate
2MWT	n/a	n/a	n/a	? Unknown	n/a	? Unknown	+ Limited
6MWT	n/a	n/a	+ Limited	++ Moderate	? Unknown	? Unknown	++ Moderate
VO_2 peak	n/a	n/a	n/a	? Unknown	? Unknown	n/a	n/a
RPE	n/a	n/a	n/a	? Unknown	? Unknown	? Unknown	n/a

3.5 Discussion

The first search of the present systematic review identified 27 outcome measures, assessing self-reported and objectively measured walking performance, self-perceived fatigue, effort of walking, QoL, balance, falls and lower limb function, that had been used in studies assessing the effects of assistive technology to treat foot drop in pwMS. The most frequently used measure was the 10mWT (n=19), followed by 3D gait kinematics (n=12) and PCI (n=10). Interestingly, although 3D gait kinematics was one of the most frequently used outcome measures to assess the effects of assistive technology to treat foot drop, its psychometric properties have not yet been reported for this specific population (Bethoux & Bennett, 2011). Similarly, there were no psychometric studies identified for PCI for the MS population. However, studies into the psychometric properties for 3D gait kinematics have demonstrated that 3D gait analysis is a reliable, valid and responsive tool for characterizing gait in stroke sufferers (Yavuzer et al, 2008), CP (Noonan et al, 2003; Kainz et al, 2017b; Nieuwenhuys et al, 2017) and many musculoskeletal disorders (Laroche et al, 2011; Bates et al, 2016). Similarly, the construct validity of the PCI has been assessed in the subacute stroke population and its reliability documented in children with cerebral palsy (Raja et al, 2007; Delussu et al, 2014).

The second, and main, search for studies assessing the psychometric properties of the 20 outcome measures related to walking performance, lower limb function and effort of walking identified in the first search, revealed 41 studies that evaluated only 10 of these twenty outcomes. Of those 10 measures, the MSWS-12 was found to have a strong level of evidence for its internal consistency and test-retest reliability and the T25FW for

construct validity. Moderate evidence was found for the test-retest reliability and responsiveness of the 6MWT and the responsiveness of the 10mWT.

Short distance walking tests, such as the 10mWT and T25FW have been classified as reliable owing to ICC values of 0.7 and over. However, there are indications that walking speed, as measured over such short distances, may not be appropriate to assess the benefits of functional electrical stimulation for community walkers with relatively low levels of disability. For example, Miller et al. (2016) found that pwMS who walked faster than 0.8m/s did not increase their walking speed in the T25FW with the assistance of functional electrical stimulation, while those with a slower walking speed than 0.8m/s did.

De Vet et al. (2006b) distinguished two aspects of reliability, namely consistency (or relative reliability), which is assessed by the ICC and secondly measurement error (or absolute reliability), which is reported by measures like SEM, MDC and the LoA. Although ICC values are informative, these are greatly dependent on inter-subject variance in the outcome measure. The knowledge of the measurement error of a particular measure is essential for both researchers and clinicians when selecting a reliable outcome as both need to establish whether an "improvement" in a patient's walking performance, with the use of assistive technology, is due to measurement error or a 'true' change as a result of the intervention (Bruton et al, 2000). This is best achieved via the implementation of MDC data, the value beyond which, in this instance, a difference between performance with and without assistive technology can be considered a true change. In our review, of the 18 studies evaluating 'relative' reliability, only four also reported the measurement error of six outcomes (MSWS-12, 10mWT, T25FW, 6MWT, RPE and VO2peak). The

MDC was reported to be 22 points, 2.7s and 88 meters for the MSWS-12, T25FW and 6MWT respectively (Learmonth et al, 2013a). Paltamaa et al. (2005) reported an SEM of 0.09 for the 10mWT that indicates an $MDC_{95\%}$ of 2.4s. Heine et al. (2015) reported an SEM of 1.1 and 0.131 for RPE and VO_2peak respectively, indicating an $MDC_{95\%}$ of 3.04 for RPE and 0.36 $L \cdot min^{-1}$ for VO_2peak. However, the strength of the results in these studies rated as 'indeterminate' because the MIC values were not reported. According to Terwee et al. (2009b) 81], the value of the measurement error needs to be considered in relation to MIC (also referred to as the MCID) values in order to determine whether the measurement error of an outcome measure is acceptable for use in research or clinical practice. If the measurement error is exceeding MIC, it is difficult to interpret whether the observed changes are clinically relevant and are not just because of measurement error (de Vet et al, 2006a; Terwee et al, 2009b). Another issue to consider is that patient-related factors, such as medications and comorbidities, can influence clinical outcome measurement findings by contributing to measurement error. Many people with MS using medications and have co-morbidities and symptoms such as fatigue, which may change over a period of several weeks or even days (Powell et al, 2017; Kasser et al, 2017). These factors are likely to affect outcome measures, both in test-retest reliability studies and clinical trials. One of the items in COSMIN checklist for reliability and measurement error is: 'Were patients stable in the interim period on the construct to be measured?' For an 'excellent' score for this item authors need to provide evidence that the patients were stable. However, none of the papers, including those with repeated assessment over more than two weeks (Learmonth et al, 2013a; Heine et al, 2015), reported this evidence.

The methodological quality of the 41 studies rated according to the COSMIN criteria revealed that both the analysis and reporting of the psychometric properties of outcome measures is often inappropriate. For example, the methodological quality of responsiveness studies was often only rated as 'fair' and 'poor' because the hypotheses were not reported or because there was a lack of information regarding the comparator instruments (often EDSS) and their psychometric properties. Another potential problematic issue with evaluating responsiveness was that in eight of the 14 studies there was no intervention and the (often assumed) hypothesis was that pwMS would deteriorate over the time frame of the study, which ranged from one to two years.

The comparator instrument in seven out of the 15 studies that evaluated validity was the EDSS, which has been widely accepted as a gold standard to measure disability in pwMS. However, its use as a gold standard to validate outcomes of walking performance may be less appropriate. The EDSS (Kurtzke, 1983) is a scale that was developed over 30 years ago and even though studies have reported high inter- and intra-rater reliability and high correlations for face validity (Goodkin et al, 1992; Sharrack et al, 1999), there are other studies raising issues regarding its reliability and objectivity and whether it can be considered a 'gold standard' (Hobart et al, 2000; Cohen et al, 2012).

It should be noted that the aforementioned methodological issues in the studies included in this review do not imply that the outcome measures are not appropriate but instead that more psychometric studies with higher methodological quality are needed. When planning studies to assess the psychometric properties of outcome measures, researchers should

consult standard guidelines such as COSMIN in relation to the selection of appropriate study design, statistical analysis and reporting of methods and results.

To our knowledge, this is the first review that evaluated the evidence for the psychometric properties of walking performance related measures used to assess the effect of assistive technology in pwMS. We used standardized criteria to evaluate both the methodological quality (COSMIN) and quality of the results (Terwee et al, 2007; Terwee et al, 2012). To date, only two reviews have tried to highlight which are the most useful tools for walking assessment in pwMS. However, one was a narrative review of available outcome measures and offered little detail about psychometric properties (Bethoux & Bennett, 2011). The other was a topical review including some details of the psychometric properties of measures to assess walking disability, but which did not employ specific criteria to evaluate the evidence for their use (Kieseier & Pozzilli, 2012). Work has been published on the stroke population that evaluated, also using COSMIN criteria, the psychometric properties of walking performance measures (van Bloemendaal et al, 2012). This review concluded that most of the outcome measures were reliable and valid for use in the stroke population, but it was observed, similar to our findings, that there was a lack of evidence for the minimally important change and responsiveness. Two COSMIN reviews into the functional outcomes in the cerebral palsy population came to similar conclusions (Amman-Reiffer et al, 2014; Zanudin et al, 2017).

This review has several limitations. Firstly, the COSMIN checklist was originally designed for patient-reported outcome measures and not for performance-based measures such as the majority of those included in our review. However, as there is no specific

checklist for performance-based measures we opted to use the COSMIN checklist since most of the items scored are also highly relevant to performance-based measures. Additional rules were specified for the ratings of items that were only applicable to patient-reported outcome measures. Moreover, another limitation might be that in our initial search we used only two interventions (i.e. AFOs and FES) to identify walking performance outcome measures, while studies reporting on other orthotic interventions may also have reported on walking outcomes. However, we believe that these two interventions were the most appropriate, since they are used widely to treat foot drop and the outcome measures used would be appropriate for that reason. Furthermore, only studies published in English, Greek or Dutch were included, which means that eligible studies in other languages will likely have been excluded. Finally, in the majority of the included studies, the mean EDSS was four or more and five studies involved participants with a mean EDSS of six. The responsiveness and reliability of walking performance measures in pwMS with EDSS > 4 may be different from those who are less affected by MS.

3.6 Conclusion

The present systematic review reported on the psychometric properties of outcome measures used to assess the effects of assistive technology to treat foot drop. Forty-one studies were identified which reported information on the psychometric properties of only 10 of the previously identified 20 measures related to walking performance. Strong levels of evidence were found for internal consistency and test-retest reliability of the MSWS-

12 and the construct validity of the T25FW. Moderate evidence was found for the test-retest reliability and responsiveness of the 6MWT and for the responsiveness of the 10mWT. None of the outcome measures that were evaluated for measurement error had an acceptable level of evidence for this measurement property. Our findings do not indicate that the existing outcome measures included in this review are poor, but that there is a need for more high quality studies evaluating the psychometric properties of these measures. Future research should (i) investigate the psychometric properties, and in particular measurement error and responsiveness, of a wider range of walking performance related measures and (ii) use standard guidelines such as the COSMIN to increase methodological quality enabling clinicians and researchers to select appropriate outcome measures to assess the effects of assistive technology to treat foot drop.

Chapter 4. Test-retest reliability and minimal detectable change of ankle kinematics and spatiotemporal parameters

4.1 Purpose of chapter

The purpose of this chapter is to describe the methodology, present and discuss the results of two studies reporting on relative and absolute reliability of the ankle kinematics and spatiotemporal parameters associated with walking in pwMS. These two studies report on the kinematics and spatiotemporal parameters of two groups of pwMS with a different level of walking impairments.

4.2 Introduction

The typical gait pattern in most pwMS is to walk slowly, with associated shorter stride length and prolonged double support phase (Benedetti et al, 1999; Givon et al, 2009; Kelleher et al, 2010; van der Linden et al, 2014b). Three-dimensional (3D) gait analysis is an established method to quantify and reveal even minimal gait disorders in a variety of populations. Increasingly, studies have incorporated 3D motion capture systems to record the gait kinematics of pwMS (e.g. Scott et al, 2013; Pau et al, 2014; McLoughlin et al, 2016; Barr et al, 2017).

Outcome measures, such as 3D gait kinematics, need to be reliable and responsive to changes, as both are essential psychometric characteristics for evaluating meaningful changes after clinical practice or research interventions (Hopkins, 2000). Variability in 3D kinematics between sessions can be either due to 'intrinsic' factors, such as age and pathology or due to 'extrinsic' factors such as marker placement, data processing or

assessors' experience. Consequently, it is important to identify the measurement error in order to avoid misinterpretation of the results, i.e. either meaningful changes to be missed or small changes to be considered meaningful (Schwartz et al, 2004; McGinley et al, 2009).

Chapter 3 reported on a recent systematic review regarding the psychometric properties of walking performance, effort of walking and lower limb function measures that have been used to evaluate assistive technology to treat foot drop in pwMS. This review included 41 studies reporting on the psychometric properties of ten outcome measures of walking performance (Andreopoulou et al, 2018). Eighteen studies investigated the reliability (intra-rater, inter-rater & test-retest) of eight outcome measures through ICC but only four studies reported on measurement error of six outcome measures by presenting SEM and MDC values. Interestingly, although 3D gait kinematics was one of the most frequent outcome measures to assess the effects of assistive technology to treat foot drop, no studies reporting on its psychometric properties were found for the MS population even though it is considered a 'gold standard' for the assessment of walking performance (Bethoux & Bennett, 2011).

Although the psychometric properties of 3D gait kinematics might not have been extensively investigated in the MS population, these psychometric properties are well established in other populations [i.e. healthy, cerebral palsy (CP), stroke, etc.]. Table 4.1 presents a summary of the studies reporting on relative and absolute reliability of ankle kinematics and stride parameters in healthy, stroke and CP populations. After a comprehensive literature search in PubMed and in relevant systematic reviews, twenty studies were identified (n=8 healthy, n=5 stroke, n=7 CP) which examined test-retest,

intra-rater and inter-session reliability through ICC values (i.e. relative reliability or 'consistency'), while a few studies also reported on the measurement error, presented as MDC values. Most of the studies reported excellent ICC values for peak dorsiflexion (DF) in swing and angle ankle at initial contact (IC) apart from two studies on healthy individuals, one study in stroke patients for the ankle angle at IC and one study in CP population with Gross Motor Function Classification System (GMFCS) level II. However, as discussed in Chapter 3, there is a lack of MS specific information on 3D gait kinematics. For that reason, it would be valuable to obtain relative and absolute reliability information for this specific group, since reliability may be dependent on the level of gait impairment and could differ among different patient populations and even among the same population with different disease progression. Furthermore, it is of importance to have reliable measures, and especially for ankle kinematics, that can objectively assess the effects of FES as a treatment for foot drop in pwMS.

Therefore, the aim of this study was to examine two aspects of reliability in two groups of pwMS with different levels of walking impairment.

• The first objective was to examine relative reliability which was assessed by the ICC.

• The second objective was to examine absolute reliability (i.e. measurement error) of the ankle kinematics and spatiotemporal parameters.

Table 4.1 Summary of studies reporting on psychometric properties of ankle kinematics and gait speed in healthy, stroke and CP populations.

Study	Participant characteristics [n, age (years), type]	Type of reliability, session number, interval	ICC	MDC
Healthy				
Fernandes et al, 2016	n = 23, age: 35	Test-retest reliability, 2 x sessions, interval: 7-11 days	Peak DF: ICC = 0.77	Peak DF: MDC = 3.9°
Kadaba et al, 1989	n = 40, age range: 18-40	Test-retest reliability, 3 x sessions, interval: 1 week	Ankle DF/PF: CMC_{left} = 0.937 CMC_{right} = 0.933	–
Kaufman et al, 2016	n = 10, age: 30	Inter-session reliability, 3 x sessions, interval: within 12 months	Medial inter-session error: 1.2°	–
Meldrum et al, 2014	n = 30, age: 30	Test-retest reliability, 2 x sessions, interval: 1 day – 1 week	Peak DF in swing: ICC = 0.44 Ankle angle at IC: ICC = 0.49 Gait speed: ICC = 0.89	Peak DF in swing: MDC = 8.09° Ankle angle at IC: MDC = 8.26° Gait speed: MDC = 0.17m/s

Mentiplay & Clark, 2018	n = 30, age: 23	Test-retest reliability, 2 x sessions, interval: 7 days	Peak DF in swing: ICC = 0.78	Peak DF in swing: MDC = 3.8°
Maynard et al, 2003	n = 10, age: 39.2	Intra-rater reliability, 2 x sessions, interval: 1 week	Ankle angle at IC: ICC = -0.10	–
Monaghan et al, 2007	n = 10, age: 28.5	Intra-rater reliability, 2 x sessions, interval: 1 week	Ankle sagittal: $ICC_{0\%stance} = 0.94$ $ICC_{50\%stance} = 0.90$ $ICC_{50\%swing} = 0.93$	–
Wilken et al, 2012	n = 29, age: 24.7	Intrarater-intersession reliability, 2 x sessions, interval: 5 days	Ankle DF: ICC = 0.82	Ankle DF: MDC = 3.66°
Stroke				
Caty et al, 2009	n = 10, age: 53.5	Inter-session reliability, 3 x sessions, interval: baseline (T_0), 1 day (T_1) & 1 month (T_2)	Peak DF in swing: $ICC_{T0-T1} = 0.84$ $ICC_{T0-T2} = 0.67$ Ankle Angle at IC: $ICC_{T0-T1} = 0.64$	–

			ICC_{T0-T2} = 0.44	
Correa et al, 2017	n = 20, age: 55.2	Test-retest reliability, 2 x sessions, interval: 2-7 days	GDI (for one stride): $ICC_{paretic}$ = 0.69 $ICC_{non-paretic}$ = 0.80	GDI (for one stride): $MDC_{paretic}$ = 9.4° $MDC_{non-paretic}$ = 7.5°
Devetak et al, 2016	n = 17, age: 54.9	Test-retest reliability, 2 x sessions, interval: 2-7 days	GPS: $ICC_{paretic}$ = 0.92 $ICC_{non-paretic}$ = 0.93 $ICC_{overall}$ = 0.95	GPS: $MDC_{paretic}$ = 2.3° $MDC_{non-paretic}$ = 1.9° $MDC_{overall}$ = 1.7°
Kesar et al, 2011	n = 19, age range: 47-75	Between-session reliability, 2 x sessions, interval: 20 days	Peak DF in swing: ICC = 0.941 Ankle angle at IC: ICC = 0.893	Peak DF in swing: MDC = 4.9° Ankle angle at IC: MDC = 7.0°
Yavuzer et al, 2008	n = 20, age: 54.2	Test-retest reliability, 2 x sessions, interval: 2 hours	Ankle DF in stance: ICC = 0.96 Ankle PF in swing: ICC = 0.95	–
Cerebral palsy				

Klejman et al, 2010	GMFCS I: n = 10, age: 6.6 GMFCS II: n = 10, age: 8.1 GMFCS III: n = 8, age: 7.3	Test-retest reliability, 2 x sessions, interval: 1-2 weeks	Peak DF in swing: $ICC_{GMFCS\ I} = 0.97$ $ICC_{GMFCS\ II} = 0.89$ $ICC_{GMFCS\ III} = 0.87$ Ankle angle at IC: $ICC_{GMFCS\ I} = 0.92$ $ICC_{GMFCSII} = 0.55$ $ICC_{GMFCS\ III} = 0.90$	Peak DF in swing: $MDC_{GMFCS\ I} = 7.8°$ $MDC_{GMFCS\ II} = 11.6°$ $MDC_{GMFCS\ III} = 11.4°$ Ankle angle at IC: $MDC_{GMFCS\ I} = 10.1°$ $MDC_{GMFCSII} = 11.6°$ $MDC_{GMFCS\ III} = 5.1°$
Mackey et al, 2005	n = 10, age: 9	Test-retest reliability, 2 x sessions, interval: 1 week	Ankle DF/PF: $CMC_{normal} = 0.98$ $CMC_{hemiplegic} = 0.96$	–
Miller et al, 1996	n = 5, children with CP n = 5, healthy children age range: 6-16	Test-retest reliability, 5 x sessions, interval: NS	Ankle DF/PF: $ICC_{CP} = 0.941$ $ICC_{healthy} = 0.598$	–
Rasmussen et al, 2015	n = 18, age: 8	Intra-rater reliability, 2 x sessions, interval: 7-10 days	GPS overall: ICC = 0.88 GDI overall: ICC = 0.88	GPS overall: MDC = 0.25° GDI overall: MDC = 10.80

Redekop et al, 2008	GMFCS I: n = 11, age: 7.0 GMFCS II: n = 12, age: 8.2 GMFCS III: n = 10, age: 10.0	Intra-session reliability, 8 x trials, interval: NS	Ankle angle at IC: $ICC_{GMFCSI} = 0.90$ $ICC_{GMFCSII} = 0.77$ $ICC_{GMFCSIII} = 0.88$ Gait speed: $ICC_{GMFCSI} = 0.76$ $ICC_{GMFCSII} = 0.83$ $ICC_{GMFCSIII} = 0.83$	Ankle angle at IC: $MDC_{GMFCSI} = 8.04°$ $MDC_{GMFCSII} = 3.71°$ $MDC_{GMFCSIII} = 4.21°$ Gait speed: $MDC_{GMFCSI} = 0.22m/s$ $MDC_{GMFCSII} = 0.22m/s$ $MDC_{GMFCSIII} = 0.28m/s$
Sorsdahl et al, 2008	n = 18, age range: 3-13	Test-retest reliability, 1 x session, interval: 25min	Cadence: ICC = 0.95	Cadence: MDC = 11.08steps/min
Steinwender et al, 2000	n = 20, children with CP n = 20, healthy children age range: 7-15	Test-retest reliability, 3 x sessions, interval: 1 week	Ankle DF/PF: $CMC_{cp} = 0.83$ $CMC_{normal} = 0.87$	–

Abbreviations: CMC: Coefficient of Multiple Correlation; CP: Cerebral Palsy; DF: Dorsiflexion; GDI: Gait Deviation Index; GMFCS: Gross Motor Function Classification System; GPS: Gait Profile Score; IC: Initial Contact; ICC: Intraclass Correlation Coefficient; MDC: Minimal Detectable Change

4.3 Methods

4.3.1 Participants

For the purpose of this study, the data of two groups of pwMS who participated in two different studies conducted at the QMU motion analysis laboratory were analyzed.

The data collected for 'group A' are described in more detail in Chapter 5. This data collection was conducted by the PhD candidate and took place between 2017-2018. The eligibility criteria for this group were clinically definite MS according to the revised McDonald criteria, stable MS which meant no evidence of disease activity based on clinical and radiological findings, have minimal disability in no more than two functional systems (EDSS < 3.5), regularly take part in aerobic exercise (for at least 30 minutes twice a week) and aged between 18 and 80. The exclusion criteria were pregnancy or breast-feeding and comorbidities such as neurological conditions other than MS, cardiovascular or respiratory diseases.

Data collection for 'group B' took place in the period 2010-2011 by another researcher and were used for secondary analysis of the data. Participants were eligible if they were diagnosed with MS according to the revised McDonald criteria, experiencing foot drop during walking, being judged suitable for FES to treat foot drop by a clinical specialist physiotherapist and aged between 18 and 75. The exclusion criteria were any relapse in the past three months and patients who were pregnant or breast-feeding.

4.3.2 Ethical opinion

Both studies gained a favourable ethical opinion from NHS (National Health Service) Research Ethics Committee and Queen Margaret University Ethics Committee and NHS

Lothian Research and Development approval was sought before commencing the studies. All procedures were in accordance with the declaration of Helsinki regarding human experimentation.

4.3.3 Recruitment process

Participants in group A were recruited through the Anne Rowling Regenerative Clinic and the MS Therapy Centre Lothian, both based in Edinburgh, UK. Potential participants were identified by the MS specialist nurses and were given the Participant Information Sheet (PIS). Once they expressed an interest to take part, a visit was arranged to the Anne Rowling Clinic to meet with a consultant neurologist and co-investigator (Dr. Don Mahad) who explained the protocol in detail and answered any questions.

All participants who were eligible and agreed to take part in the studies signed an informed consent form prior to commencing with the protocol.

Participants for group B were recruited through a community NHS physiotherapy service based in Edinburgh, UK. People with MS who were considered suitable by their clinician for FES treatment were invited to take part in the study by giving them a PIS. People who agreed to take part in the study after reading the PIS, were contacted by the principal investigator (PI) to arrange the visits to QMU.

4.3.4 Study procedures

Both groups visited the motion analysis laboratory at Queen Margaret University and underwent 3D gait analysis assessment. Since the two groups involved participants taking

part in discrete studies, about seven years apart, the procedure of measurements for both groups is described separately.

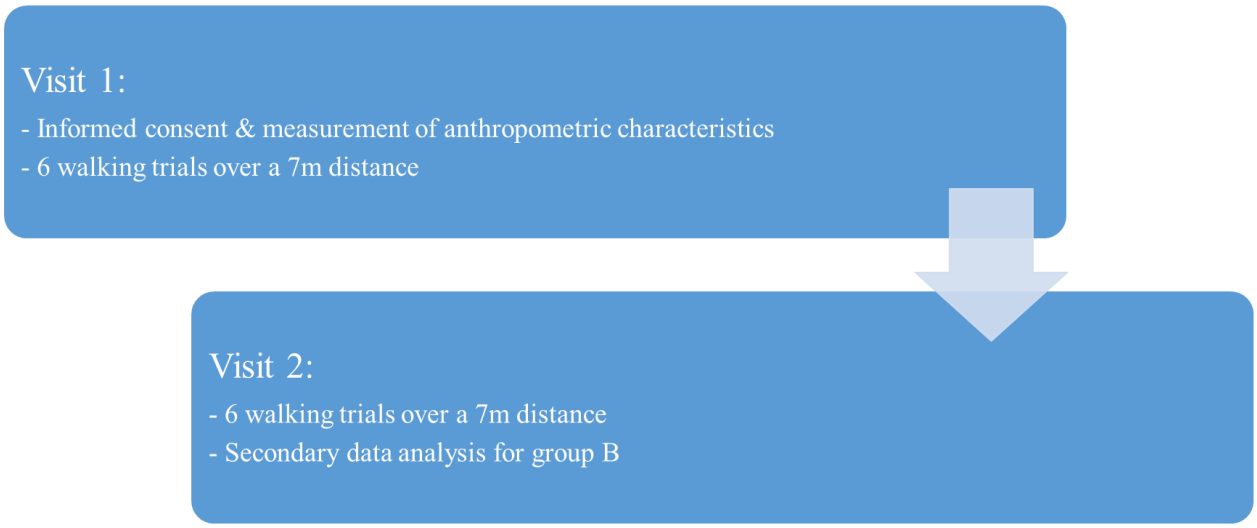


Figure 4.1 Flow diagram of the study design of Chapter 4.

4.3.4.1 Group A

Participants visited the laboratory on three occasions. The first visit consisted of measuring the anthropometric characteristics of the participants, recording of baseline gait kinematics during overground walking and habituation on the treadmill for the exercise task that they would have to complete in the second and third visit. The assessments in the second and third visits were identical. Walking kinematics of the participants were recorded before and after a 20 minute jogging/running on a treadmill. The protocol that this group completed is described in more detail in Chapter 5. For the purpose of this chapter, we used the kinematic variables derived from the gait analysis before the exercise

task from the second and third visit to calculate the test-retest reliability and the kinematic variables before the exercise task from the second visit for the intra-session reliability.

4.3.4.2 Group B

Participants were asked to visit the laboratory at least four times within a maximum of six weeks. All visits were separated by at least three days, but no more than 14 days. In each visit, participants underwent gait analysis assessment, along with the 6-min walk test and the 10-m timed walking test, the latter test was performed twice in each assessment. The 10-m walk test and the gait analysis assessment were carried out with and without FES, within one session, while the 6-min walk test was performed either with or without FES in one session. The 6-min walk test and 10-m timed walking test were not used in the analysis for the present study. Gait analysis was carried out by first conducting the trials without FES, followed by the trials with FES.

As previously reported (see section 4.3.1), this was a secondary analysis of the data collected in a previous study (Scott et al, 2013). For the purposes of the present study, i.e. to estimate the test-retest reliability, we used the data from the trials without FES from the first and second visit and for estimating intra-session reliability we used the trials without FES from the first visit.

4.3.5 Gait analysis protocol (both studies)

The anthropometric characteristics that were measured were: body mass, height, leg length, knee width, ankle width and tibial torsion. These measurements are required for the calculation of the 3D position of the ankle, knee and hip joint centers by the Vicon Plug-in-Gait (Vicon Motion Systems, Oxford, UK). The gait analysis was undertaken using a 100Hz eight infra-red camera Vicon Nexus (version 1.8.5) computerized 3D motion capture system (Vicon Motion Systems, Oxford, UK). Passive reflective sphere markers of 9 or 14mm were placed on the lower limbs and pelvis of the participants according the Helen Hayes marker system (Kadaba et al, 1990). Markers were placed on the anterior superior iliac spines (ASIS) of each side and one on middle between the two posterior superior iliac spines (PSIS). For each leg, markers were placed on the lateral side of the thigh and tibia (wand markers), the knee marker on the lateral epicondyle of the femur just above the joint line in the center of the knee (walking trials only), the toe marker on the point between the second and third metatarsal heads, a marker on the middle of the lateral malleolus and the heel marker on the calcaneum in vertical alignment with the toe marker (Figure 4.2). A static trial (with the participant standing) was conducted first using a knee alignment device (KAD) to derive the orientation of the knee flexion/extension axis. After

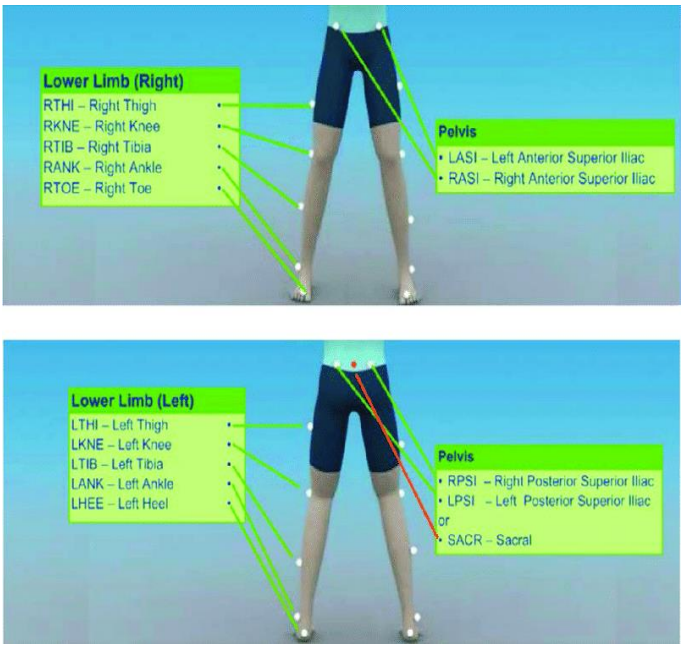


Figure 4.2 Anterior and posterior view of marker placement according to Helen Hayes marker system.

the static trial and the removal of the KADs, a standard reflective marker was placed on the lateral side of the knee as described above for the walking trials. Participants underwent gait analysis barefoot and were instructed to walk a distance of 7m across the laboratory, which consisted one trial, at their preferred self-selected walking pace. At least six trials were recorded using the Vicon motion analysis system. Participants were able to rest between trials if wished to do so. Similar to standard practice in clinical gait analysis, we chose the number of six walking trials to take into account variability and this number also allows trials to be discarded due to possible issues such as missing markers (i.e. not visible by at last two cameras).

4.3.6 Data analysis

Pelvis and lower limb angles in sagittal, transverse and frontal planes were derived using the Vicon Plug-in-Gait software. Kinematic data for each of the six trials in each visit were time normalized so that every trial included one gait cycle (i.e. between two consecutive foot strikes) consisting of 51 data points. Polygon software (version 3.5.2) (Oxford Metrics Group, Oxford, UK) was used for reporting and graphically presenting the kinematic data for each trial. Through Polygon, kinematic and spatiotemporal data were extracted to Microsoft Excel files and then a custom-made program in Matlab R2014b (Mathworks, Natrick, USA) was used for further analysis. The peak dorsiflexion (DF) in the swing phase and the ankle angle at initial contact (IC) were derived for each of the six walking trials together with the spatiotemporal parameters (walking speed, step length, cadence). The average value of the joint angles (peak DF in swing and the ankle angle at IC) and step length was calculated from the six walking trials for the most and

the least affected legs for each of the two visits. The total walking speed and cadence were derived for both legs for each visit.

The Gait Profile Score (GPS) (Baker et al, 2009) was also calculated for each walking trial for both visits and for each left and right leg separately and were summed to derive the total GPS score. The GPS is an index of overall gait pathology and is derived from the pelvis, hip, knee and ankle kinematics. The GPS in this study was calculated from the Gait Variable Scores (GVS) of the sagittal ankle angle, the foot progression angle, the sagittal knee angle, sagittal and frontal plane hip angle and pelvic angles in all planes. We excluded the hip rotation from the GPS score, as it has been documented that this angle is very sensitive to even small changes in marker placement and therefore exhibits large measurement error (Schwartz et al, 2004). The GVS is the root mean square (RMS) differences over the whole gait cycle of an individual with MS and the average data of a healthy control group of similar age range. The healthy group consisted of eleven individuals (5 male, 6 female) and with an average age of 49.5 years (van der Linden et al, 2014b). The higher the GPS score, the higher the deviation from a normal gait pattern.

4.3.6 Statistical analysis

4.3.6.1 Test-retest reliability

4.3.6.1.1 Consistency

The relative reliability for peak DF in swing, ankle angle at IC, GPS and spatiotemporal parameters was calculated with the ICC (model 2,2) using a two-way mixed effects type of average measures for absolute agreement. We chose this type of ICC because repeated measurements (test-retest) cannot be regarded as randomised samples (Koo & Li, 2016).

Intraclass correlation coefficient values ≥ 0.75 were as regarded as excellent level of practical and clinical significance for test-retest reliability, while ‘good’ was between 0.60-0.74 and ‘fair’ between 0.40-0.59 (Shrout & Fleiss, 1979; Cicchetti, 1994).

4.3.6.1.2 Agreement

The absolute reliability for the aforementioned variables was determined by reporting the SEM, the MDC and the LoA.

The SEM is related to an outcomes’ reliability, because it provides an indication of the variability among measurements (de Vet et al, 2006 (b)). It was determined with the following equation:

$$\text{SEM} = SD \times \sqrt{(1 - ICC)} \qquad (1)$$

where SD is the standard deviation from the first testing session.

Minimal detectable change is important information for an evaluative instrument to provide to clinicians and researchers because it gives information of the cut-off point above which it can be regarded as ‘true’ change (de Vet et al, 2006) and was calculated using the equation:

$$\text{MDC}_{95\%} = 1.96 \times SEM \times \sqrt{2} \qquad (2)$$

Limits of agreement were quantified in this chapter for both groups by the Bland-Altman (B&A) plot. The B&A plot was developed in order to describe the agreement between two different instruments measuring the same outcome or the repeatability (precision) of

one instrument on repeated occasions (Bland & Altman, 1999). In order to do so, firstly by subtracting the mean of the first visit from the second visit the mean difference was calculated. The LoA were calculated according to the equation:

$$\text{LoA} = Mean \pm (SD \times 1.96) \qquad (3)$$

where, Mean is the average of the difference of the two visits and SD is the standard deviation of the difference between the two visits.

4.3.6.2 Intra-session reliability

4.3.6.2.1 Consistency

The relative intra-session reliability (i.e. over the six trials) for peak DF in swing, ankle angle at IC, GPS and spatiotemporal parameters was calculated with the ICC (model 2,2) using a two-way mixed effects type for average measures for absolute agreement. The ICC was calculated from the six walking trials of the first visit for both groups. As mentioned before, ICC values ≥ 0.75 were regarded as excellent, while 'good' were between 0.60-0.74 and 'fair' between 0.40-0.59 (Shrout & Fleiss, 1979; Cicchetti, 1994).

4.3.6.2.2 Agreement

The absolute reliability for peak DF in swing, ankle angle at IC, GPS and spatiotemporal parameters was determined by calculating the SEM and $MDC_{95\%}$ for these variables. For estimating the SEM and $MDC_{95\%}$ values, the equations 1 and 2 were used respectively. The standard deviation of the first walking trial was used in the equations.

All mean, standard deviation (SD) and ICC values and B&A plots were calculated using SPSS 23 (IBM, Armonk, USA) and SEM and $MDC_{95\%}$ calculations were determined with the use of Microsoft Excel 2007 (Microsoft Corp., Redmond, WA).

4.4 Results

Group A consisted of 21 highly active pwMS, with the data being collected by the researcher. Group B consisted of 28 pwMS who presented and were being treated for foot drop. This was a secondary analysis of previously collected data. The descriptive characteristics of the participants in group A and B can be seen in Table 4.2.

Table 4.2 Participant characteristics of both groups.

	Group A	**Group B**	**p-value**
Female/Male, n	14/7	14/14	0.12
Age, years	43.8 (10.9)	52.2 (10.1)	**0.004**
EDSS range	1-3.5	4-6	-
Height, m	1.71 (0.08)	1.69 (0.07)	0.27

Body mass, kg	72.6 (14.4)	78.2 (16.7)	0.10
BMI, kg/m^2	24.7 (4.4)	26.9 (4.3)	**0.04**
Walking aid (none/ walking stick/stroller), n	21/0/0	16/11/1	-

Abbreviations: BMI: Body Mass Index; EDSS: Expanded Disability Status Scale; NA: not available

4.4.1 Test-retest reliability

4.4.1.1 Consistency

Table 4.3 and Table 4.4 present the mean and standard deviation values from both visits of both groups respectively , ICCs [95% Confidence Intervals (CI)] between the two visits, and the SEM and $MDC_{95\%}$ for all kinematic and spatiotemporal parameters. For group A all kinematic and spatiotemporal parameters exhibited 'excellent' ICC values. For group B, DF in swing, ankle angle at IC, walking speed, step length and cadence all exhibited 'excellent' ICC values of ≥ 0.75. 'Good' reliability was shown by the ICC values for the GPS of the most affected leg, while ICC values for GPS of the least affected leg indicated 'fair' reliability.

4.4.1.2 Agreement

The agreement analysis showed that group B had higher SEM values than group A by ≈1° and higher $MDC_{95\%}$ values by approximately ≈1°-5° for all the kinematic variables. Group

A also had lower SEM and $MDC_{95\%}$ values for walking speed, step length and cadence (Table 4.3 & 4.4).

Table 4.3 Test-retest reliability with mean (SD), ICC (95%CI), SEM and $MDC_{95\%}$ for group A kinematic and spatiotemporal parameters.

	Session 1 Mean (SD)	**Session 2 Mean (SD)**	**p-value**	**ICC (95% CI)**	**SEM**	**$MDC_{95\%}$**
DFMA (°)	7.9 (2.3)	8.1 (3.0)	0.30	0.862 (0.660-0.944)	0.85	2.4
DFLA (°)	9.0 (2.5)	8.9 (2.8)	0.40	0.862 (0.657-0.944)	0.9	2.5
ICMA (°)	0.9 (4.2)	1.7 (4.6)	0.06	0.919 (0.800-0.967)	1.21	3.4
IC LA (°)	2.5 (3.9)	2.1 (5.2)	0.27	0.852 (0.635-0.940)	1.5	4.2
GPSMA (°)	9.0 (1.4)	8.7 (1.1)	0.12	0.698 (0.274-0.876)	0.75	2.1
GPSLA (°)	9.0 (1.1)	8.9 (1.6)	0.45	0.621 (0.039-0.848)	0.7	1.9
WS (m/s)	1.25 (0.14)	1.26 (0.18)	0.36	0.833 (0.585-0.932)	0.06	0.16
SLMA (m)	0.64 (0.07)	0.64 (0.08)	0.31	0.931 (0.830-0.972)	0.02	0.05
SLLA(m)	0.64 (0.07)	0.63 (0.07)	0.24	0.909 (0.777-0.963)	0.02	0.06
Cadence (steps/min)	117 (7)	119 (9)	0.053	0.877 (0.698-0.950)	2	6

Abbreviations: DFMA: peak dorsiflexion in swing of the most affected leg; DFLA: peak dorsiflexion in swing of the least affected leg; GPSMA: Gait Profile Score of the most affected leg; GPSLA: Gait Profile

Score of the least affected leg; ICMA: ankle angle at initial contact of the most affected leg; ICLA: ankle angle at initial contact of the least affected leg; ICC: Intraclass Correlation Coefficient; MA: most affected leg; MDC: minimal detectable change; LA: least affected leg; SEM: standard error of measurement; SLMA: step length of the most affected leg; SLLA: step length of the least affected leg; WS: walking speed.

Table 4.4 Test-retest reliability with mean (SD), ICC (95% CI), SEM and $MDC_{95\%}$ for group B kinematic and spatiotemporal parameters.

	Session 1 Mean (SD)	**Session 2 Mean (SD)**	**p-value**	**ICC (95% CI)**	**SEM**	**$MDC_{95\%}$**
DFMA (°)	2.9 (6.7)	1.9 (6.5)	0.12	0.891 (0.761-0.951)	2.2	6.1
DFLA (°)	7.5 (4.2)	6.4 (3.6)	**0.04**	0.823 (0.606-0.920)	1.8	4.9
ICMA (°)	-3.7 (6.9)	-4.5 (5.6)	0.20	0.840 (0.647-0.928)	2.8	7.7
ICLA (°)	1.8 (5.6)	1.1 (5.0)	0.23	0.773 (0.495-0.898)	2.7	7.4
GPSMA(°)	9.1 (1.3)	8.8 (1.0)	0.18	0.636 (0.192-0.836)	0.8	2.2
GPSLA (°)	9.5 (1.1)	9.2 (0.8)	0.09	0.489 (-0.105-0.767)	0.8	2.2
WS (m/s)	0.77 (0.21)	0.81 (0.2)	0.10	0.831 (0.629-0.924)	0.08	0.23
SLMA(m)	0.49 (0.09)	0.51 (0.09)	0.09	0.848 (0.666-0.931)	0.04	0.1
SLLA (m)	0.47 (0.08)	0.49 (0.10)	0.15	0.742 (0.431-0.884)	0.04	0.1
Cadence (steps/min)	93 (16)	96 (14)	**0.04**	0.927 (0.835-0.968)	4	12

Abbreviations: DFMA: peak dorsiflexion in swing of the most affected leg; DFLA: peak dorsiflexion in swing of the least affected leg; GPSMA: Gait Profile Score of the most affected leg; GPSLA: Gait Profile Score of the least affected leg; ICMA: ankle angle at initial contact of the most affected leg; ICLA: ankle angle at initial contact of the least affected leg; ICC: Intraclass Correlation Coefficient; MA: most affected leg; MDC: minimal detectable change; LA: least affected leg; SEM: standard error of measurement; SLMA: step length of the most affected leg; SLLA: step length of the least affected leg; WS: walking speed.

Limits of Agreement were for the two visits for both groups and the Bland-Altman plots were examined and are presented in Figures 4.2-4.6 for group A and Figures 4.7-4.11 for group B for all kinematic and spatiotemporal parameters.

Figures 4.3 to 4.7 show the limits of agreement for group A and indicates that the limits of agreement for group A were narrower than those for group B for all kinematic and spatiotemporal parameters.

For peak DF in swing (Figure 4.3) the mean difference was very close to zero with one outlier in both most and least affected legs.

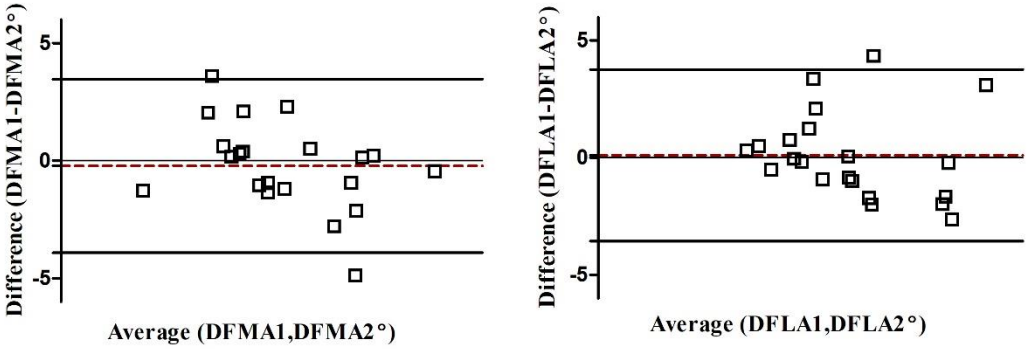


Figure 4.3 Bland & Altman plot of peak DF in swing (most and least affected leg) for group A.

─ 95% Limits of Agreement
-- Mean difference
DFMA: Peak dorsiflexion in swing of the most affected limb; DFLA: Peak dorsiflexion in swing of the least affected limb

The limits of agreement of IC for the least affected leg were wider (-6.1-7.1°) than the most affected leg (-5.4-3.8°), but the participants were clustered around the mean difference compared to the most affected leg where they were more dispersed between the mean ± 1.96 x SD.

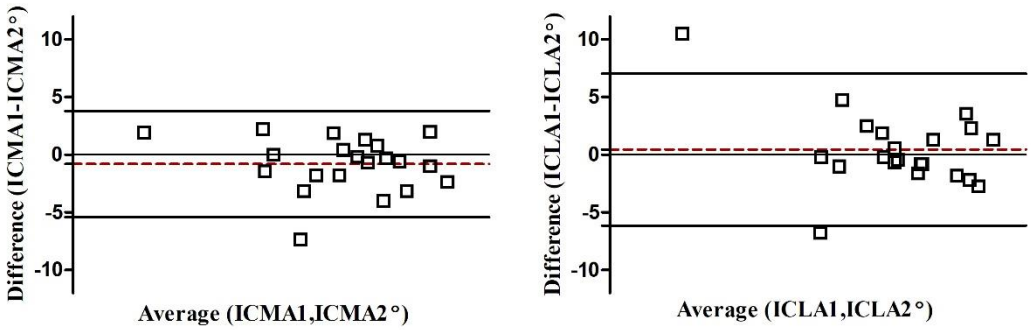


Figure 4.4 Bland & Altman plot of IC (most and least affected leg) for group A.

─ 95% Limits of Agreement
-- Mean difference
ICMA: Ankle angle at initial contact of the most affected limb; ICLA: Ankle angle at initial contact of the least affected limb

Similar to the ankle kinematics, the limits of agreement for the GPS in group A were narrower (≈-2.5-2.5°) compared to group B.

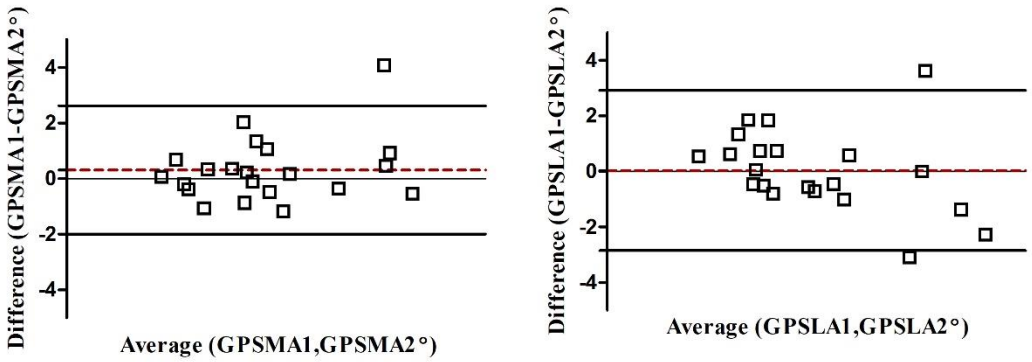


Figure 4.5 Bland & Altman plot of GPS (most and least affected leg) for group A.

─ 95% Limits of Agreement
-- Mean difference
GPSMA: Gait profile score of the most affected limb; GPSLA: Gait profile score of the least affected limb

For both walking speed and cadence all values were between the limits of agreement apart from one outlier in each. The LoA were similar for walking speed in both groups, but smaller for cadence in group A.

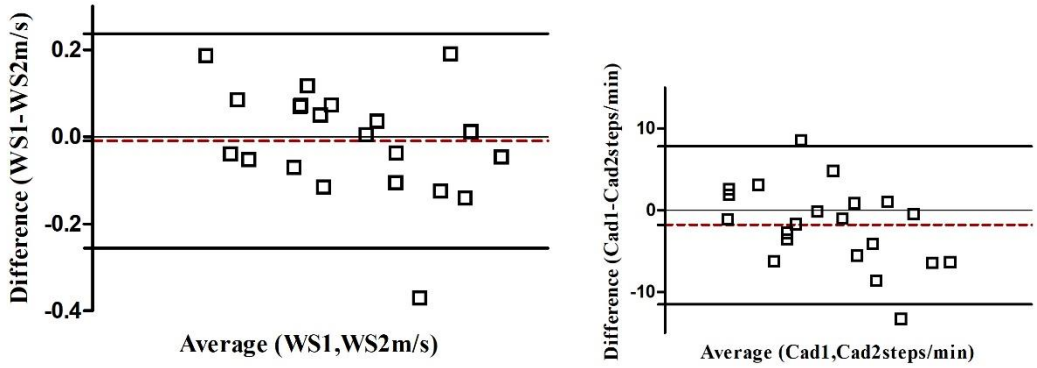


Figure 4.6 Bland & Altman plot of walking speed and cadence for group A.

─ 95% Limits of Agreement
-- Mean difference
WS: Walking speed

There was a small negative deviation away from zero for step length in both legs, meaning that step length was higher in the second visit. A similar trend of dispersion was observed for both sides.

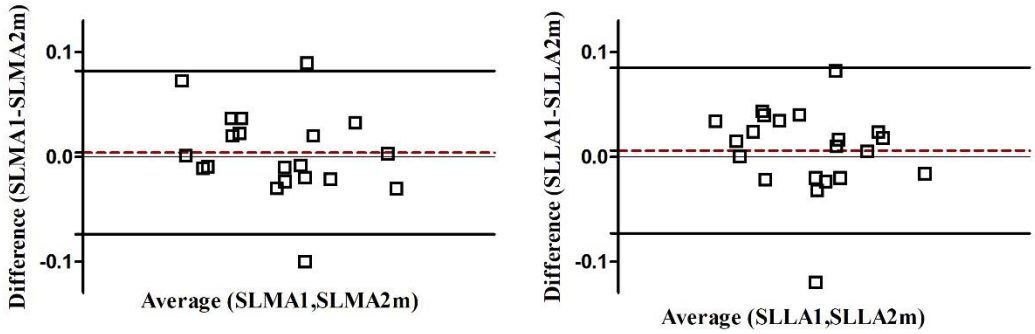


Figure 4.7 Bland & Altman plot of step length (most and least affected leg) for group A.

─ 95% Limits of Agreement
-- Mean difference
SLMA: Step length of the most affected limb; SLLA: Step length of the least affected limb

Figure 4.8 to Figure 4.12 presents the limits of agreement for group B. Figure 4.8 shows that there are two outliers for peak DF in swing (i.e. difference between the visits of more than 1.96 x SD). It can also be seen that the limits of agreement are wider for the most affected leg (-7.1-9.0°) than for the least affected leg (-4.7-6.8°).

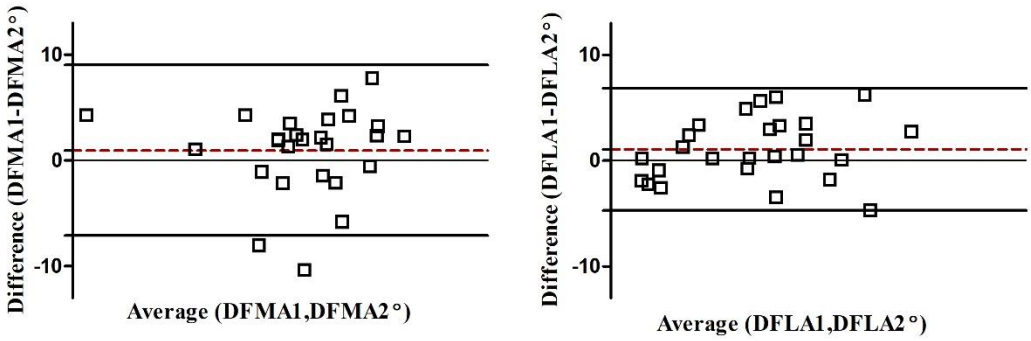


Figure 4.8 Bland & Altman plot of peak DF in swing (most and least affected leg) for group B.

─ 95% Limits of Agreement
-- Mean difference
DFMA: Peak dorsiflexion in swing of the most affected limb; DFLA: Peak dorsiflexion in swing of the least affected limb

For the ankle angle at IC there was the same trend as for DF in swing. For the least affected leg, all participants were concentrated around the mean difference of just above zero apart from two outliers. For the most affected leg, values were more dispersed but within the limits of agreement apart from one outlier. Limits of agreement were similar for both legs.

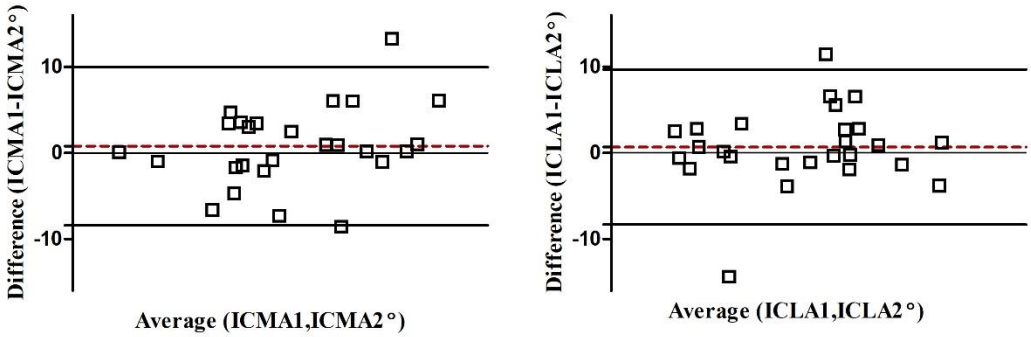


Figure 4.9 Bland & Altman plot of IC (most and least affected leg) for group B.

─ 95% Limits of Agreement

-- Mean difference
ICMA: Ankle angle at initial contact of the most affected limb; ICLA: Ankle angle at initial contact of the least affected limb

The limits of agreement for the GPS of both most and least affected legs were narrow (less than 3°).

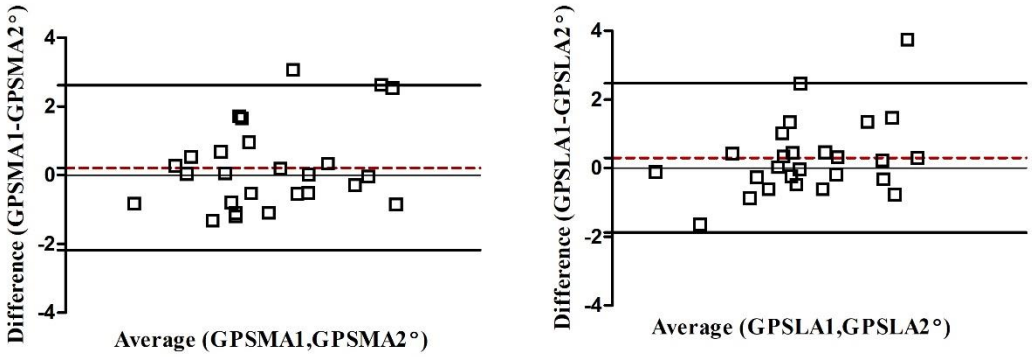


Figure 4.10 Bland & Altman plot of GPS (most and least affected leg) for group B.

─ 95% Limits of Agreement
-- Mean difference
GPSMA: Gait profile score of the most affected limb; GPSLA: Gait profile score of the least affected limb

For walking speed and cadence a trend was observed with participants walking slightly faster and with more steps/min in the first compared to the second visit. For both parameters, there was a central tendency around the mean with only one outlier for both parameters.

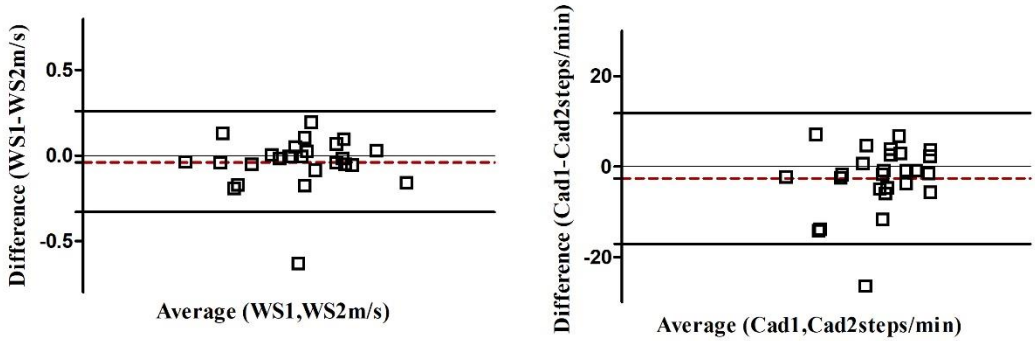


Figure 4.11 Bland & Altman plot of walking speed and cadence for group B.

─ 95% Limits of Agreement
-- Mean difference
WS: Walking speed

The limits of agreement for step length were narrow (≈-0.15-0.12m) with only two and one outliers for the most and least affected legs respectively.

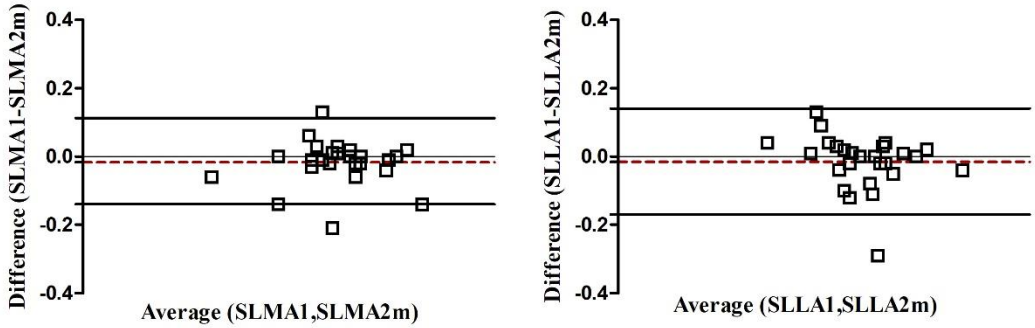


Figure 4.12 Bland & Altman plot of step length (most and least affected leg) for group B.

─ 95% Limits of Agreement
-- Mean difference
SLMA: Step length of the most affected limb; SLLA: Step length of the least affected limb

4.4.2 Intra-session reliability

4.4.2.1 Consistency

The intra-session reliability (i.e. over six trials) was derived by calculating the ICC for the most and least affected legs for peak DF in swing, ankle angle at IC, GPS and spatiotemporal parameters. Both groups presented excellent consistency, with ICC values > 0.90 for all the kinematic and spatiotemporal variables (Table 4.5 & 4.6).

4.4.2.2 Agreement

For group A SEM values for peak DF in swing, ankle angle at IC and GPS were ranging between 0.1°-1.3°, with MDC values ranging from 0.3°-3.7° and were lower than what group B presented. The same pattern was for spatiotemporal parameters too, with group A presenting lower SEM and $MDC_{95\%}$ values for walking speed, step length and cadence.

Table 4.5 Intra-session reliability with mean (SD), ICC, 95% CI, SEM, $MDC_{95\%}$ for Group A for the kinematic and spatiotemporal parameters.

	Mean (SD)	**ICC (95% CI)**	**SEM**	**$MDC_{95\%}$**
DFMA (°)	7.8 (2.5)	0.949 (0.905-0.977)	0.7	1.9
DFLA (°)	9.0 (2.6)	0.976 (0.954-0.989)	0.5	1.3
ICMA(°)	1.2 (4.2)	0.986 (0.972-0.994)	0.6	1.5
ICLA(°)	2.5 (4.4)	0.945 (0.897-0.975)	1.3	3.7
GPSMA(°)	9.1 (1.4)	0.992 (0.986-0.997)	0.1	0.3

GPSLA(°)	8.9 (1.3)	0.980 (0.963-0.991)	0.2	0.5
WS (m/s)	1.3 (0.2)	0.982 (0.965-0.992)	0.02	0.06
SLMA (m)	0.64 (0.08)	0.982 (0.966-0.992)	0.01	0.03
SLLA(m)	0.64 (0.07)	0.984 (0.971-0.993)	0.01	0.02
Cadence (steps/min)	118 (8)	0.975 (0.953-0.989)	1	3

Abbreviations: DFMA: peak dorsiflexion in swing of the most affected leg; DFLA: peak dorsiflexion in swing of the least affected leg; GPSMA: Gait Profile Score of the most affected leg; GPSLA: Gait Profile Score of the least affected leg; ICMA: ankle angle at initial contact of the most affected leg; ICLA: ankle angle at initial contact of the least affected leg; ICC: Intraclass Correlation Coefficient; MA: most affected leg; MDC: minimal detectable change; LA: least affected leg; SEM: standard error of measurement; SLMA: step length of the most affected leg; SLLA: step length of the least affected leg; WS: walking speed.

Table 4.6 Intra-session reliability with mean (SD), ICC (95% CI), SEM, $MDC_{95\%}$ for group B for the kinematic and spatiotemporal parameters.

	Mean (SD)	**ICC (95% CI)**	**SEM**	**$MDC_{95\%}$**
DFMA (°)	2.9 (6.9)	0.989 (0.980-0.995)	0.7	1.9
DFLA (°)	7.4 (4.4)	0.985 (0.972-0.993)	0.5	1.5
ICMA (°)	-3.7 (7.7)	0.961 (0.929-0.981)	1.4	4.0
ICLA (°)	1.7 (6.3)	0.950 (0.910-0.976)	1.6	4.4
GPSMA (°)	9.1 (1.4)	0.954 (0.916-0.978)	0.3	0.9
GPSLA (°)	9.6 (1.3)	0.949 (0.909-0.975)	0.3	0.8

WS (m/s)	0.77 (0.2)	0.983 (0.969-0.991)	0.03	0.08
SLMA(m)	0.50 (0.1)	0.961 (0.930-0.982)	0.02	0.05
SLLA (m)	0.49 (0.1)	0.977 (0.956-0.989)	0.01	0.03
Cadence (steps/min)	89 (23)	0.995 (0.991-0.998)	2	5

Abbreviations: DFMA: peak dorsiflexion in swing of the most affected leg; DFLA: peak dorsiflexion in swing of the least affected leg; GPSMA: Gait Profile Score of the most affected leg; GPSLA: Gait Profile Score of the least affected leg; ICMA: ankle angle at initial contact of the most affected leg; ICLA: ankle angle at initial contact of the least affected leg; ICC: Intraclass Correlation Coefficient; MA: most affected leg; MDC: minimal detectable change; LA: least affected leg; SEM: standard error of measurement; SLMA: step length of the most affected leg; SLLA: step length of the least affected leg; WS: walking speed.

All kinematic variables in group B had higher SEM and $MDC_{95\%}$ values than group A. All peak DF in swing and ankle angle at IC SEM values for group B were between 0.5°-1.6° and GPS of 0.3°, with $MDC_{95\%}$ values ranging from 1.74°- 4.31° and ≈ 0.80° respectively.

4.5 Discussion

The present study aimed to report on relative and absolute reliability of 3D ankle kinematics and spatiotemporal parameters in two groups of pwMS with different levels of walking impairment. These values were calculated through a test-retest design with a

period of seven to fourteen days between the two visits and also an intra-session design by using the six walking trials of the first visit.

In Chapter 3 the gap of reliability indices for 3D gait kinematics in the MS population was highlighted, even though it is considered a 'gold standard' and is increasingly used to assess the effects of FES. This is the first study to report on such indices (i.e. ICC, SEM, MDC) for gait kinematics and spatiotemporal parameters in the MS population.

Reliability is considered to be the extent to which the scores of an outcome on repeated measurements of participants that have not change in the interim period are the same under several conditions (e.g. over time, by different assessors, etc.) (de Vet et al, 2006). Relative reliability (i.e. consistency) in the present study was reported through ICC for a test-retest design and also during the same session (intra-session). Absolute reliability (i.e. agreement) was reported through SEM, $MDC_{95\%}$ and Bland & Altman plots.

4.5.1 Consistency

Intra-class Correlations Coefficients were calculated for both groups for test-retest reliability and indicated good to excellent reliability for peak DF in swing, ankle angle at IC, GPS, walking speed, step length and cadence in both groups. The only exception to this was the GPS for group B which showed fair ICC values (≈0.48-0.63). The reliability of 3D gait kinematic and kinetic data have been explored in many studies in the healthy population. Similar to our findings, studies with healthy participants have shown good to excellent ICC values ranging from 0.77-0.93 for peak DF in swing and ankle angle at IC (Kadaba et al, 1989; Monaghan et al, 2007; Wilken et al, 2012; Fernandes et al, 2016; Mentiplay & Clark, 2018). However, other studies reported low ICC values for peak DF

in swing at ankle angle at IC (Meynard et al, 2003; Meldrum et al, 2014). Both of these studies reporting low ICCs (≈-0.10-0.44) had a small sample size and high standard deviation of each outcome, indicating inter-participant variability.

For the spatiotemporal parameters, such as cadence, step length, walking speed and step width, we observed good to excellent ICC values in both groups. These findings are in accordance with studies in healthy populations which reported that these spatiotemporal parameters are reliable and highly repeatable (Meldrum et al, 2014). Although estimates of reliability for gait kinematics in pwMS have not been reported, a study by Sosnoff et al. (2015) examined the reliability of walking speed, cadence and step length. They reported that in a group of MS people with a varied disability level [Patient Determined Disease Steps (PDDS) range 0-6] there were excellent ICC values (0.91) for the spatiotemporal parameters which is similar to our findings for both groups.

Reliability of kinematic variables has extensively been examined in the stroke population with studies reporting good to excellent ICC values indicating high test-retest reliability for peak DF in swing, ankle angle at IC and GPS (Yavuzer et al, 2008; Kesar et al, 2011; Devetak et al, 2016). On the contrary, one study indicated lower ICC values for most of the kinematic variables and especially for the ankle angle at heel strike (Caty et al, 2009). All studies presented in Table 4.1 with stroke population had a small sample size ranging from ten to twenty participants, with some studies reporting reliability values only for the paretic limb. Devetak et al (2016) examined ICC values for both limbs and found that the paretic limb had lower ICC values than the non-paretic.

In children with cerebral palsy, sagittal kinematic parameters were found to have good to excellent reliability (Steiwender et al, 2000; Mackey et al, 2005; Redekop et al, 2008).

When analysing the reliability for the GMFCS level separately, relative reliability in all kinematic variables was highest for children with GMFCS Level I (least impaired walking ability), with ICC values for GMFCS Level II and III ranging from moderate to high (≈0.55-0.90) (Redekop et al, 2008; Klejman et al, 2010). The reliability of cadence, walking speed, stride length, step length and single stance time were found to be excellent (ICC range 0.76-0.95) (Redekop et al, 2008; Sorsdahl et al, 2008), while step width had low ICC values indicating poor reliability (Sorsdahl et al, 2008).

Intra-session reliability was also examined in the present study and we found excellent ICC values (> 0.92) for all kinematic and spatiotemporal parameters for both MS groups. There are few studies reporting on intra-session reliability, but all of them are in agreement with our findings showing excellent ICC values in different population groups (Steinwender et al, 2000; Monaghan et al, 2007; Yavuzer et al, 2008; Redekop et al, 2008; Kesar et al, 2011; Devetak et al, 20016). These findings of all studies, including ours, was expected and can be explained by the fact that compared to test-retest reliability in intra-session reliability designs there is a decrease in the number of both extrinsic and intrinsic factors that could affect reliability such as marker placement and possible change in participants such as day-to-day variability especially in patient populations.

4.5.2 Agreement

Absolute agreement in the present study was examined through the indices of SEM, $MDC_{95\%}$ and Limits of Agreement. For the test-retest design, the $MDC_{95\%}$ values of peak DF in swing and ankle angle at IC were low (≈2.5º) for the low EDSS group (group A), while for the high EDSS group (group B) SEM and $MDC_{95\%}$ values were higher (4.9º-

7.7°). The same trend was also observed for GPS, walking speed, cadence and step length when comparing the two groups. The $MDC_{95\%}$ values for our low EDSS group (group A) were similar to the values reported for the healthy population. Three studies with healthy participants reported MDC of ≈3.8° for peak DF in swing (Wilken et al, 2012; Fernandes et al, 2016; Mentiplay & Clark, 2018). However, Meldrum et al (2014) found high $MDC_{95\%}$ values for both peak DF in swing and ankle angle at IC of around 8°.

Studies reporting on SEM and MDC values on kinematic and spatiotemporal parameters on stroke population and children with CP have similar findings as those in our group with higher EDSS (group B). The MDC values for stroke population have been reported to be of ≈2° for the GPS, 4.9° for peak DF in swing and 7° for initial contact (Kesar et al, 2011; Devetak et al, 2016). In children with CP, the higher the GMFCS level, the higher the SEM and MDC values for all kinematic and spatiotemporal parameters (Redekop et al, 2008; Sorsdahl et al, 2008; Klejman et al, 2010; Rasmussen et al, 2015). Hence, the results of our study and studies with different patient populations suggest that the more the walking pattern is impaired, the less reliable the measurement of gait kinematics will be. The association between gait variability and clinical walking indices in the MS population has been explored previously and it was found that people with higher EDSS (> 4.5) and using assistive devices had great variability in spatiotemporal parameters (i.e. step length, step time, etc.) than people with lower EDSS (Socie et al, 2013; Socie et al, 2014; Kalron, 2016).

Limits of agreement in our study were narrower for the minimally disabled group compared to the group with higher EDSS, which can be indicated by the increased variability in people with more impaired walking. Only two studies with healthy

population have reported on the limits of agreement for peak DF in swing, walking speed and cadence (Meldrum et al, 2014; Fernandes et al, 2016) and the results were similar to the minimally disabled group of MS. The limits of agreement should be interpreted with caution by clinicians, since there are no other MS or different patient group population data available to compare.

The indices of SEM and $MDC_{95\%}$ for the intra-session design were lower for both MS groups than for the test-retest design, which is in accordance with a study in stroke population (Kaser et al, 2011). There are not many studies examining the intra-session SEM and MDC values and that could potentially be because it is expected to have lower values than a test-retest reliability design since the number of extrinsic factors that could affect the repeatability is decreased.

4.5.3 Limitations

This study has some limitations that should be addressed in the future. Firstly, both groups, but especially group A, had small sample sizes. According to the COSMIN criteria the methodological quality of a study would be considered poor, if the sample size is less than 30.

Marker placement and analysis of the kinematics in the two groups were performed by two different researchers. This could introduce a form of an extrinsic error in the results when comparing the reliability indices between the two groups, i.e. the researcher placing the markers in group B may have had higher test-retest marker placement error, explaining somewhat the lower reliability for this group. However, both researchers for the two groups followed the same training protocol conducted by the same Director of the Gait

Lab, and similar number of practice sessions, including the assessment of within day test-retest reliability of gait kinematic parameters for healthy participants.

4.6 Conclusion

The main objective of the present study was to determine relative and absolute reliability of 3D ankle kinematics and spatiotemporal parameters in two groups of pwMS with different levels of walking impairment.

The results showed good to excellent ICC values of peak DF in swing, ankle angle at IC, GPS, walking speed, step length and cadence. The SEM and $MDC_{95\%}$ values for each of the parameters were lower for the group with lower EDSS compared to the group with higher EDSS and that suggests that the higher the walking impairment the lower the inter-session reliability.

The findings of this study provide clinicians and researchers with the indices of relative and absolute reliability for ankle kinematics in pwMS which can be applied to both clinical decision making and in the design of studies aimed at treating foot drop in people with MS.

Chapter 5. Gait characteristics in minimally impaired people with MS after an exercise task

5.1 Purpose of the chapter

The purpose of this chapter is to describe the methodology, and to present and discuss the findings of a study that was conducted to investigate the exercise-induced gait deterioration (as a measure of fatigability) in a highly physically active group of pwMS. The second aim of this study was to investigate whether the exercise-induced gait deterioration occurs in a healthy control group.

5.2 Introduction

As described in more detail in Chapter 2, gait impairments are a common symptom in pwMS with approximately 85% reporting experiencing walking difficulties at some stage throughout the disease course (Kelleher et al, 2010; Bethoux & Bennett, 2011). Walking impairments in relapsing remitting MS might be because of a relapse, which will frequently recover by leaving minimal or no permanent neurological deficit.

Anecdotal reports in clinical practice indicate that a subset of pwMS, who are regularly engaged in exercise (i.e. jogging, running, etc.), often experience a transient phenomenon that often manifests as foot drop that resolves after cessation of exercise or a short period afterwards and the ability to walk and perform physical activities, such as running, returns to baseline. It is thought that this exercise induced transient phenomenon could be a sign of fatigability. At present, these transient periods of performance deterioration, or

fatigability, during exercise is largely put down to a rise in core body temperature in MS, which is known as Uhthoff's phenomenon. It is considered as part of a temperature related neurological dysfunction when myelin or the protective fatty sheaths of nerve cells is damaged or lost (demyelinated) (Frohman et al, 2013). However, studies linking real-time core temperature changes with this reversible phenomenon are lacking.

In order to explore how common it is for pwMS to experience transient foot drop, we conducted an audit. It was found that 33 out of 47 MS respondents (70%) reported experiencing this phenomenon with three reporting permanent foot drop and 11 not reporting any foot drop. The audit also showed that the most commonly activities associated with the self-reported foot drop were walking and running, and with foot drop appearing to be more evident and occurring earlier during faster speeds or when walking uphill (see details in section 2.3.3 and Appendix 1).

Existing literature has demonstrated the presence of fatigability captured by gait kinematic changes. These studies have used protocols of prolonged walking until exertion or shorter walking tests (i.e. 6MWT) and reported exercise induced gait changes, evidenced as decrease in dorsiflexion at initial contact and at peak dorsiflexion in swing (Sehle et al, 2011; McLoughlin et al, 2016; van der Linden et al, 2018). However, the EDSS of the MS population in these studies varied greatly ranged from 1.0 to 6.0, with the majority 4.0 or above (i.e. walking ability is affected) (McLoughlin et al, 2016; van der Linden et al, 2018). To date, there has been no evidence whether exercise-induced deterioration occurs in pwMS with minimal neurological disability (EDSS < 3.0).

Therefore, the aim of this study was to investigate the presence of exercise-induced deterioration of gait kinematics, thought to be sign of fatigability, in a group of pwMS

who have stable relapsing remitting MS, presented with minimal neurological disability (EDSS < 3.0).

• The first objective was to investigate the differences in ankle kinematics and spatiotemporal parameters before and directly after a 20-minute treadmill run at a self-selected speed and we hypothesised that the fatigability evidenced as foot drop was present in this group of pwMS after the exercise task.

• The second objective was to find out whether this transient foot drop would be evident in an age-matched healthy control group and to explore the differences between the (unfatigued) gait kinematics of this group of pwMS and healthy controls. We hypothesised that the healthy control group would not exhibit this transient foot drop after the exercise task.

• The third objective was to investigate the differences in exercise-induced gait deterioration between the second and third visit on the group of pwMS. It was hypothesised that there would be a similar pattern of gait deterioration in the group of pwMS between the two visits.

• Finally, the fourth objective was to explore the association between the presence of exercise-induced deterioration of gait kinematics and changes in core temperature in a subgroup of pwMS.

5.3 Methods

5.3.1 Design

This was a pilot observational study with a pre-post design to measure exercise-induced gait deterioration, evidenced as foot drop in pwMS. The principal investigator (PI) for this study was a consultant neurologist from the Anne Rowling Clinic, Dr. Mahad, whose work in the laboratory using genetically manipulated mice with neuron specific mitochondrial respiratory chain enzyme defects indicated the role of neuronal metabolic changes in motor fatigability. Based on his laboratory work, this study was designed to investigate motor fatigability in the MS population. Data collection for this study took place from March 2017 until December 2018.

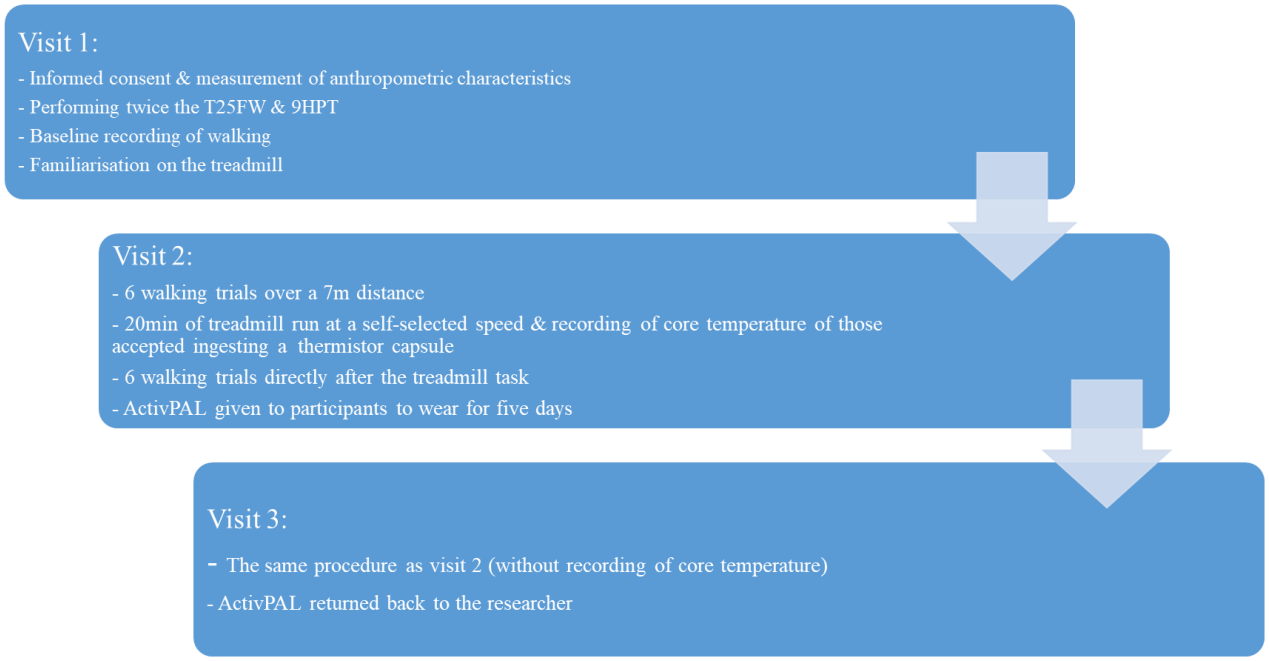


Figure 5.1 Flow diagram of study design for Chapter 5.

5.3.2 Participants

Eligibility criteria required participants to be 18 years and over, have confirmed diagnosis of MS according the revised MacDonald criteria (Polman et al, 2011) and minimal

disability in no more than two functional systems (EDSS < 3). Further, in order to be eligible for inclusion, participants had to be regularly engaging in exercise for at least 30 minutes continuously twice a week and be able to jog or run a mile without stopping. Only participants with stable MS were included in the study. Stable MS was defined as no evidence of disease activity based on clinical and radiological evidence within a two-year period prior to enrollment. Participants were eligible whether or not they were prescribed MS disease modifying therapy (DMT).

The exclusion criteria were comorbidities such as neurological conditions other than MS, cardiovascular disease, respiratory disease such as asthma, metabolic disorders such as diabetes and thyroid disease, and peripheral vascular disease. In addition, people with a history of temperature sensitivity unrelated to exercise and those who were pregnant or post-partum were excluded from the study.

For the healthy control group, eligibility criteria also required participants to be physically active (regularly engaged in exercise for at least 30min twice a week or more) and aged above 18. The exclusion criteria were similar to those in the MS group.

Finally, contraindications for ingesting the thermistor capsule (VitalSense Core Tempereature Capsule, Equivital Inc., Cambridge) (see section 5.3.6.2) were: a) weight less than 36.3 Kg, b) diagnosed with diverticulitis, inflammatory bowel disease, gag reflex disorders, previously gastrointestinal (GI) surgery or hypo-mobility of the GI tract, c) undergo MRI while the thermistor capsule is still on the body, and d) having a cardiac pacemaker or implanted electronic medical device. All these contraindications are recommended by the manufacturer (Equivital Inc., Cambridge).

5.3.3 Ethical opinion

Favourable ethical opinion was obtained from the National Health Service (NHS) Research Ethics Committee (South East Scotland Ethics Committee 02, REC reference number: 15-SS-0088) (Appendix 6) and Queen Margaret University Ethics Committee. Further, NHS Lothian Research and Development approval was also obtained before commencing with the study (Appendix 7). All procedures were in accordance with the declaration of Helsinki with regards to human participation. Participants could withdraw from the study at any point without giving any reason.

5.3.4 Recruitment process

People with MS with minimal neurological disability were recruited through the Anne Rowling Clinic (ARC), Edinburgh. The ARC is an outpatient MS service, where the patients are seen once a year by a neurologist consultant and every three to four months by MS specialist nurses. Patients are routinely asked about their lifestyle, including exercise, during their consultations. MS nurses and consultants mentioned the study to potentially eligible participants and if interested provided them with a PIS (Appendix 8). Potential participants were informed that their decision with regard to taking part in the study would not affect their usual medical care. Potential participants were given two or more weeks to read the PIS and decide whether or not they were willing to take part in the study. Those that were willing to take part were invited to arrange a phone call or a visit to ARC to discuss the study with the PI and have the opportunity to ask questions in order to finalize their eligibility. Following the verbal consent to the PI, the first visit to the

motion analysis laboratory was arranged by the student where written informed consent (Appendix 9) was obtained after further explanation of the study.

Healthy age matched control individuals were recruited through QMU and ARC by word of mouth. Those healthy individuals that expressed an interest were given the PIS and if they were willing to take part, a visit to QMU was arranged. The healthy control group also signed informed consent form prior to commencing the study protocol.

5.3.5 Study protocol

The study protocol required the participants with MS to visit the motion analysis laboratory at QMU on three occasions with one to three weeks in between each visit. Prior to arrival for the first visit, participants were sent four questionnaires by post and were asked to complete these and return them back to the student on the day of their visit. The questionnaires that were included were the Physical Activity Readiness Questionnaire for Everyone (PAR-Q+), FSMC, the FSS and a Foot Drop questionnaire developed for an audit (see details in section 5.3.6.5).

5.3.5.1 Visit one

Prior to commencing with the procedures, participants were explained the study aims and the tasks that they were going to be asked to perform.

The aim of the first visit was for the participants to become familiar with the laboratory setting and exercising on a treadmill and for baseline 3D gait analysis. The first visit lasted no more than an hour.

The first step in the process was to perform the Nine-Hole Peg Test (9-HPT) one time with each hand and the T25FW that was performed twice. The next step was to measure anthropometric characteristics that are essential for the calculation of gait kinematics and these include body height and mass, leg length, knee and ankle width and tibial torsion. After the placement of the reflective markers (9mm) for the recording of the gait kinematics according to the Helen Hays marker set (Kadaba et al, 1990), participants were asked to stand still for five seconds with the KAD placed on the both knees so that a static reference trial would be captured. For more details about marker placement see details in section 4.3.5. After removing the KADs and placing the knee markers, a baseline gait analysis assessment followed which included six trials of barefoot walking over distance of around 7m at their usual comfortable walking pace. Prior to gait analysis, participants had a few walking trials to familiarise themselves with barefoot walking and markers attached to their pelvis and lower limbs.

The final task for the first visit was for participants to become habituated to running on a treadmill. For this and any further visits to the gait laboratory that involved running on the treadmill, participants had to wear a harness (Wingman Harness, USA) that was attached to the ceiling in order to prevent any tripping or falling. When they felt ready the treadmill started at a low speed (0.5m/h) and the investigator increased the speed until the participants felt that they were running at their usual speed and which they could sustain for 20 minutes. This procedure lasted for 10 to 20 minutes depending on each individual and his/her confidence on running on a treadmill.

For those participants who had agreed to ingest the thermistor capsule in the second visit an appointment was arranged to hand over the thermistor so that participants could ingest it the night before this visit.

5.3.5.2 Visit 2

Participants were asked to attend the QMU motion analysis laboratory a second time about a week after their first visit. In the second visit, the same gait assessment as described for the first visit was performed. However, in the second and third visit this assessment was performed before and directly after an exercise task (see below).

Those participants who had ingested the thermistor capsule were given a chest belt to wear under their t-shirts to which the Equivital^{TM} EQ02 LifeMonitor was attached in order to record the core body temperature data. Participants had to ingest the thermistor capsule the night before if their assessment was in the morning or in the morning if their assessment was in the afternoon.

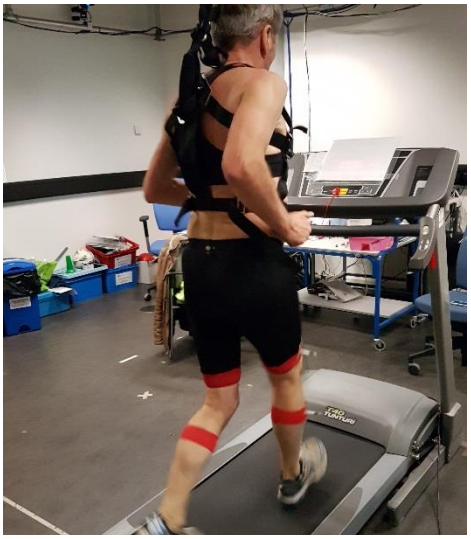


Figure 5.2 Participant performing the exercise task on the treadmill.

The exercise task comprised of jogging/running on the treadmill for 20 minutes at a self-selected speed that was perceived as their usual running speed (Figure 5.2). Participants wore a harness for safety reasons for the duration of the treadmill running task and were given five minutes for warm up and finding a speed that they were comfortable with prior to the 20-minute running task. They were informed that they could adjust the speed

throughout the task and any changes in speed and the time this occurred were recorded. Participants were allowed to stop the test at any point during these 20 minutes, although they were encouraged to continue for as long as they could. Throughout the exercise task participants were asked to rate their perceived exertion, affective valence and perceived fatigue levels every four minutes. For rating the perceived exertion the Borg Scale was used which is a 15-point scale. Affective valence (pleasure/displeasure one feels) was rated through the Feeling Scale and perceived fatigue through a Visual Analogue Scale ranging from 0 (no fatigue at all) to 10 (maximum fatigue) (see details on these scales on section 5.3.6).

Immediately after the completion of the 20-minute self-selected speed run, participants performed six walking trials during which the kinematics were recorded. As these trials required barefoot walking, participants removed their shoes and reflective markers were attached to their feet in a procedure that did not last more than three minutes. The position of the foot markers had been marked during the initial positioning, in order to assure that the markers would be placed in the same position after the removal of the shoes.

At the end of the second visit participants were given a triaxial ActivPal^{TM} activity monitor (PAL Technologies Ltd, Glasgow, UK) (see details in section 5.3.6.6) and were asked to wear this monitor for five days in order to record their daily physical activity. Participants were provided with written instructions and double sided skin friendly attachment gels (PAL-stickies^{TM}, PAL Technologies Ltd, Glasgow, UK) to attach the monitor to the front of their thigh. The instructions included details and pictures of how to attach the monitor (Appendix 10). Participants were allowed to remove the monitor at nighttime but were asked to note this on a form that was included in the instructions. The activity monitors

were returned to the student at the third and final visit to the gait lab or were posted back to the student in a stamped addressed envelope that was provided to them.

5.3.5.3 Visit 3

The final visit was mostly performed seven to ten days after the second visit or in a few cases later if participants could not attend earlier. In this visit, participants performed the same procedures as in the second visit. The same order of the tasks took place, with walking assessment followed by the exercise task and the walking assessment at the end. However, participants did not ingest the thermistor capsule again nor were asked to wear the activity monitors.

5.3.6 Outcome measures

5.3.6.1 Gait kinematics and spatiotemporal parameters

Three-dimensional (3D) gait analysis is a widely used method for identification of even subtle changes on a patients' walking pattern and is considered a 'gold' standard in terms of quantitative gait analysis. In the laboratory of QMU a 100Hz eight infra-red camera Vicon Nexus (version 1.8.5) computerized 3D motion capture system (Vicon Motion Systems, Oxford, UK) is installed and for details on this system, marker placement and gait analysis protocol please refer in Chapter 4 (section 4.3.5 and 4.3.6). Many studies have confirmed appropriate psychometric properties of 3D gait analysis, but as it was highlighted in the systematic review of Chapter 3, there was a lack of psychometric studies for 3D gait analysis in the MS population. In Chapter 4 we investigated the test-retest

reliability and the minimal detectable change of 3D ankle kinematics in two groups of pwMS with different levels of walking abilities.

For each of the six walks before and after the exercise task, one gait cycle per walk was used for further analysis. Gait kinematics and spatiotemporal parameters were derived using the Vicon Plug-in-Gait software (Vicon Motion Systems, Oxford, UK) and a custom-written Matlab script. The average of the parameters of the six gait cycles for the most and least affected leg in the MS group and the left for the healthy control group was used for analysis. The most and least affected leg of the pwMS was decided after inspection of exercise-induced change in kinematics for each leg.

The GPS (Baker et al, 2009) was also calculated for each walking trial for before and after the exercise task for each most and least affected leg separately. The GPS is an index of overall gait pathology and is derived from the pelvis, hip, knee and ankle kinematics and for more details of how this was calculated please refer to section 4.3.6.

5.3.6.2 Core body temperature

Core body temperature and how it increases during exercise in real-time is important information in MS research because of the potentially negative effect of the Uhthoff's phenomenon on activities of daily life. The EQ02 Life Monitor records multi-parameter physiological data such as heart rate, respiratory rate, skin temperature and core temperature. We were interested in the core body temperature during the 20min of the exercise task. The use of VitalSense Core Temperature Capsule transmits every 15s real-time data to the EQ02 Life Monitor. All the parameters recorded by the EQ02 Life Monitor have been examined for their psychometric properties in healthy population and

it was concluded that the monitor is a reliable and valid tool to record multi-parameter physiological data and in particular core body temperature data (Liu et al, 2013).

The core body temperature data were used to address the fourth objective, which was to examine the association between the presence of exercise-induced gait deterioration and changes in core temperature.

5.3.6.3 Perceived exertion

The Borg Scale of Perceived Exertion (Borg RPE) is an indicator of an individual's perception of the physical effort during exercise. It is a relative scale with range from 6 to 20, with 6 being 'no exerted at all' and 20 'very, very hard' (Borg, 1982) (Appendix 11). The Borg RPE scale is a widely used measure for prescribing and monitoring exercise intensity and has been found to be a reliable and valid measure in many populations such as MS, stroke and Parkinson's disease. (Sage et al, 2013; Cleland et al, 2016; Penko et al, 2017).

In the present study, perceived exertion was recorded during the 20-minute exercise task at four-minute intervals and was used as a descriptive measure of the perceived intensity which participants the participants were exercising.

5.3.6.4 Affective valence

The Feeling Scale (FS) is used to measure the affective valence during exercise (i.e. pleasure-displeasure) (Appendix 12). It is an 11-point scale which ranges from very good (+5), neutral (0) to very bad (-5). Individuals are asked to rate how they feel at specific time points, since mood can fluctuate throughout an exercise task (Hardy & Rejeski et al,

1989). As for perceived exertion, affective valence was recorded during the exercise task as a descriptor of how participants felt throughout the task.

5.3.6.5 Perceived fatigue

5.3.6.5.1 Fatigue Scale for Motor and Cognitive Functions (FSMC)

The FSMC is a 5-point Likert-type scale assessing motor and cognitive fatigue and consists from 20 items ranging from 1 (does not apply at all) to 5 (applies completely) (Penner et al, 2009) (Appendix 13). It is divided in three sub-scales, with the total score presenting cut-off values of ≥ 43 as mild fatigue, ≥ 53 as moderate fatigue and ≥ 63 as severe fatigue. For the cognitive sub-scale, a cut-off ≥ 22 represents mild cognitive fatigue, ≥ 28 moderate fatigue and ≥ 34 severe cognitive fatigue. The third sub-scale for physical fatigue presents cut-off values of ≥ 22 as mild motor fatigue, ≥ 27 as moderate motor fatigue and ≥ 32 as severe motor fatigue. The sensitivity and specificity scores indicate good internal consistency and the scale has shown good convergent and discriminant validity (Penner et al, 2009). It has also been examined for cross-cultural validity in the Danish language, indicating that is a tool that can be used in the Danish MS population (Oervik et al, 2017).

5.3.6.5.2 Fatigue Severity Scale (FSS)

The FSS is a questionnaire evaluating the impact of fatigue, primarily on the physical domain, originally developed for MS and systemic lupus erythematosus. It consists of nine items and ranges from 1 (strong disagreement) to 7 (strong agreement) reflecting the severity of fatigue the past week (Krupp et al, 1989) (Appendix 14). The total score can

range from 9-63 and a score of less than the cut-off point of 36 indicates that the person might not suffer from fatigue. The psychometric properties of the FSS has been examined for the MS population and has been found to have good test-retest reliability, precision and strong association with other outcome measures indicating good construct validity (Krupp et al, 1989; Learmonth et al, 2013b) along with cross-cultural validity (Armutlu et al, 2007; Rietberg et al, 2010).

Both the FSMC and the FSS were assessed in the present study in order to describe the level of self-reported fatigue.

5.3.6.5.3 Visual Analogue Scale to evaluate fatigue severity (VAS-F)

The VAS-F is a scale consisting of 18 items regarding subjective experiences of fatigue (Lee et al, 1991). In our study, we used the fourth item on the scale to evaluate fatigue at a specific point in time. Participants are asked to mark in a 10cm line ranging from 0 (not fatigued at all) to 10 (extremely fatigued). The scale with all the items has been found to be reliable in healthy individuals and people with sleep disorders (Lee et al, 1990) and also in the MS population (Kos et al, 2017). The VAS-F was recorded only during the exercise task, as perceived exertion and affective valence, in order to describe the overall self-reported experience of the participants during the treadmill task.

5.3.6.6 Objective habitual physical activity

ActivPAL^{TM} activity monitors provide a measure of objective habitual daily physical activity of an individual. It is attached in the front of the mid-thigh and it is small in size (35mm x 53mm x 7mm), which constitutes it unobtrusive. Participants wore the

ActivPAL™ during waking hours except while showering and swimming for five days. The accelerometer senses limb position and acceleration activity, which translates in the time an individual spends standing, lying, transitions from sit to stand and step count. Moreover, the cadence of upright activities is recorded enabling to quantify the intensity of physical activities. The psychometric properties of ActivPAL™ might not have been examined in the MS population, but have been examined in healthy (Dowd et al, 2012; Harrington et al, 2012) and other patient populations such as CP and stroke (Dahele et al, 2007; Tang et al, 2013) indicating that is a reliable and valid tool to measure physical activity.

Participants were asked to wear the ActivPAL™ over five days and was used as a descriptive measure of their habitual physical activity. We reported the daily step count, but unfortunately we were not able to correct for wear time. Although a diary to record the time, during which the ActivPAL™ was worn, was provided to the participants, a diary to record the time that they were wearing the, the majority of the participants did not record this information in the diary.

5.3.6.7 Timed 25-Foot walk

The T25FW is a short distance walking test assessing mobility function and it is one of the components of the Multiple Sclerosis Functional Composite (MSFC). The individual is instructed to walk a 25 foot distance as quickly as possible but safely twice. The score is the average of the two trials. The MSFC and the T25FW on its own have been extensively examined for their psychometric properties. The T25FW has shown to be a reliable, valid and responsive tool to examine walking capacity in the MS population

(Learmonth et al, 2012; Baert et al, 2014; Larson et al, 2013; Learmonth et al, 2013a; Andreopoulou et al, 2018).

5.3.6.8 Nine Hole Peg test

The 9HPT is a measure of manual dexterity and both dominant and non-dominant hands are tested. Individuals have to place and then remove nine pegs into nine holes as quickly as possible and the total time to finish the task is recorded. It is also one of the components of the MSFC. A recent systematic review summarized all the psychometric evidence of the 9HPT on the MS population. It is thought to be a 'gold' standard on measuring upper limb function and manual dexterity, since it has been found to be a reliable, valid, responsive and sensitive to changes outcome (Feys et al, 2017).

5.3.6.9 Thermal perception

A VAS scale for the perception of how heat is affecting participants when exercising was used. Individuals had to rate in a line from 0 to 10 how much they perceived that heat was affecting their exercise activities in their everyday life. Previous research has demonstrated good reproducibility and validity of the VAS scale for measuring thermal perception in healthy population (Davey et al, 2007; Leon et al, 2008).

The T25FW, the 9HPT and the VAS for thermal perception were recorded only during the initial visit of the participants and were used only as descriptive characteristics of the participants, in order to gain more information about their abilities.

5.3.7 Statistical analysis

In order to determine whether there was a difference between the pre- and post-exercise gait kinematic and spatiotemporal parameters in the MS group a paired t-test was carried out. The assumption of normality was checked and confirmed by visual inspection of the q-q plots and box plots of the data. A Shapiro-Wilks test was also performed, indicating that there was no violation in this assumption.

Further, an independent t-test was carried out to analyse the difference between the change in the gait kinematic and spatiotemporal parameters (i.e. before-after the exercise task) between the MS and the healthy control group. This test was also employed to examine whether differences exist between the gait parameters between the groups before the exercise task. The assumption of normality was examined as described previously within groups and the homogeneity of variance was checked using the Levene's test. There were no violations of these assumptions. The gait parameters that were analysed for both the MS and the healthy control group were the peak DF in swing, the ankle angle at initial contact, walking speed, cadence and step length.

In order to analyse the presence of exercise-induced fatigability at an individual level, we also displayed the individual change values of peak DF in swing in relation to the $MDC_{95\%}$. The $MDC_{95\%}$ values used in this study were derived from the test-retest reliability study that is described in Chapter 4.

The other outcome measures, such as the fatigue questionnaires, the T25FW and 9HPT, the ActivPAL^{TM} data and the reported scales during the exercise task, were used as descriptive measures to characterise the population of interest.

5.4 Results

5.4.1 Participants

A total of 17 pwMS with minimal disability (EDSS < 3) and who were physically active as set out in the inclusion criteria were recruited for this study. Two people withdrew from the study after their initial visit to QMU laboratory. The reasons for not continuing was personal issues for one and for the other not having enough time to commit to the subsequent visits. Table 5.1 presents the demographic characteristics of the 15 pwMS that completed the study. Further, 15 healthy individuals were also recruited and completed one visit to QMU with their demographic details also shown in Table 5.1.

Table 5.1 Demographic characteristics for MS and healthy population presented as means (SD).

	MS runners (n=15)	**Healthy control (n=15)**
Female/Male, n	9/6	8/7
Age, years	42.0 (10.0)	41.8 (11.9)
EDSS range	0.5-2.5	-
RR/PP/SP, n	15/0/0	-
Disease duration, years	16.5 (11.9)	-
T25FW, s	3.9 (1.0)	-
9HPT (Dominant/Non-Dominant), s	23.3/25.5	-

Daily step count (averaged over 5 days), n	11799 (4206) range: 3962-18832	-
$FSMC_{tot}$ (20-100)	45.5 (18.1)	-
$FSMC_{cognitive}$ (10-50)	22.3 (9.9)	-
$FSMC_{physical}$ (10-50)	23.1 (8.9)	-
FSS (9-63)	28.6 (15.7)	-

Abbreviations: 9HPT: 9 Hole Peg Test; EDSS: Expanded Disability Status Scale; FSMC: Fatigue Scale for Cognitive and Motor Function; FSS: Fatigue Severity Scale; T25FW: Timed 25 Foot Walk; RR: Relapsing Remitting; PP: Primary Progressive; SP: Secondary Progressive

The two groups were gender and age matched, with similar female to male ratio and average age in years. All the pwMS were diagnosed with relapsing remitting MS, with EDSS range of 0.5 to 2.5 and disease duration (i.e. first symptoms before confirmed diagnosis) of 16.5 years. All participants in the MS group (except one who twice a week was attending circuit gym classes) were frequent runners (at least twice a week). The healthy control group consisted of people who they ran at least three or more times a week. When inspecting the individual fatigue scores according to the cut-off values of the FSMC total score, two participants were experiencing mild fatigue, one participant moderate fatigue and four severe fatigue. Further, for the cognitive subscale of the FSMC three participants were experiencing moderate fatigue and three severe cognitive fatigue, whilst for the physical subscale one was experiencing mild physical fatigue, two participants reported moderate physical fatigue and four were experiencing severe physical fatigue. The individual scores of the FSS demonstrated that six of the participants were suffering from fatigue according to the cut-off values of the FSS [FSS < 36 (Krupp et al, 1989)].

The time interval between visits was approximately between to seven to fourteen days, with exception of three participants with interval from second to third visit been three to four weeks due to personal reasons (i.e. planned holidays and work).

5.4.2 Pre-post exercise task gait changes in pwMS in the third visit

Our first objective was to explore the differences between gait characteristics in pwMS before and directly after a 20-minute bout of treadmill running at a self-selected speed. There was a statistically significant difference between pre and post exercise peak DF in swing of the most affected limb of -1.5° [t(14)=2.703, p=0.017, 95% CI (0.32 2.76)] (Table 5.2). There were no other statistically significant differences in any other kinematic or spatiotemporal parameters between pre and post exercise walking trials (Table 5.2).

Table 5.2 Gait kinematic and spatiotemporal characteristics of the MS group before and after a 20-minute bout of treadmill running at a self-selected speed in the 3rd visit [Mean (SD)].

	Pre-run	**Post-run**	**p-value**
Kinematic parameters			
DFMA (°)	7.7 (2.9)	6.2 (3.8)	**0.017**
DFLA (°)	8.7 (2.7)	8.6 (2.7)	0.766
ICMA (°)	0.4 (4.6)	0.2 (5.1)	0.624
ICLA (°)	1.4 (5.6)	2.3 (5.8)	0.105
GPSMA (°)	8.55 (1.1)	8.53 (1.0)	0.895
GPSLA (°)	8.8 (1.4)	9.1 (1.3)	0.431
Spatiotemporal parameters			

WS (m/s)	1.29 (0.2)	1.31 (0.2)	0.447
Cadence (steps/min)	118 (10)	119 (9)	0.324
SLMA (m)	0.66 (0.08)	0.67 (0.09)	0.279
SLLA (m)	0.65 (0.07)	0.65 (0.08)	0.849

Abbreviations: DFMA: peak dorsiflexion in swing of most affected limb; DFLA: peak dorsiflexion in swing of least affected limb; GPSMA: gait profile score of most affected limb; GPSLA: gait profile score of least affected limb; ICMA: ankle angle at initial contact of most affected limb; ICLA: ankle angle at initial contact of least affected limb; SLMA: step length of most affected limb; SLLA: step length of least affected limb; WS: walking speed.

At an individual level, six out of the 15 participants with MS showed a deterioration in peak DF in swing of the most affected limb that exceeded the $MDC_{95\%}$ value (2.4°) (see Chapter 4) after a 20-minute of treadmill run at a self-selected speed (Figure 5.3). In Figure 5.3 it is apparent that no one in the healthy control group presented any deterioration in peak DF in swing exceeding the $MDC_{95\%}$ value, which could indicate that the deterioration after exercise in gait kinematics (i.e. evidenced as foot drop) is disease specific fatigability.

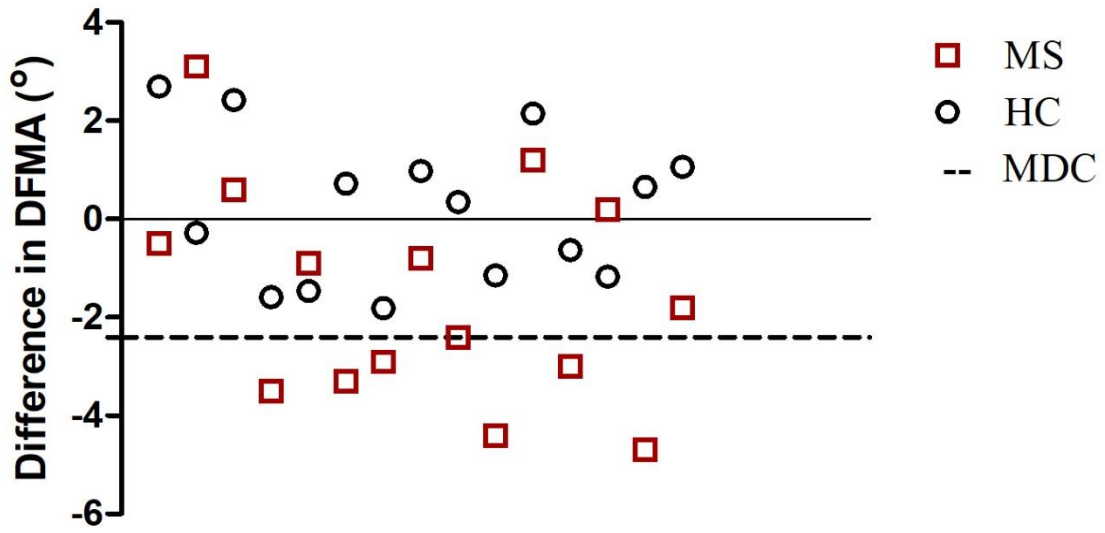


Figure 5.3 Pre-post exercise change in peak DF in swing of the MS (3rd visit) and healthy control group. Data are presented in relation to the $MDC_{95\%}$ value (see Chapter 4).

5.4.3 Gait differences between MS and healthy control group

The table below illustrates the exercise characteristics of the third visit for the MS group and for the one visit of the healthy control group (Table 5.3). The average running speed during the 20-minute treadmill run was slower for the MS group compared to the healthy control group. Although these data are descriptive, they suggest that the healthy control group exerted themselves more than the MS group as denoted by the RPE scale and reached greater levels of fatigue.

Table 5.3 Basic description of the exercise task and self-reported response characteristics for the two groups.

	MS (n=15)	**HC (n=15)**
Average Running Speed (m/s)	2.27	2.72
Peak RPE (6-20)	14	17
ΔRPE (6-20)	4.6	11
ΔFeeling Scale (+5 to -5)	6.0	5.0
ΔVAS Fatigue (0-10)	3.0	7.0

Abbreviations: RPE: Rating of Perceived Exertion; VAS: Visual Analogue Scale.

Table 5.4 presents the comparison between pre and post exercise gait characteristics in the healthy control group. As it can be seen from the table, unlike the MS group, there were no statistically significant differences in any of the kinematic parameters. Interestingly, there was a statistically significant increase after exercise in walking speed [t(14)=-3.552, p=0.003, 95% CI (-0.08 -0.02)], cadence [t(14)=-3.822, p=0.002, 95% CI (-4.3 -1.21)] and step length of the left limb [t(14)=-2.368, p=0.033, 95% CI (-0.02 -0.001)].

Table 5.4 Gait kinematic and spatiotemporal characteristics of healthy individuals before and after a 20-minute bout of treadmill running at a self-selected speed [Mean (SD)].

	Pre-run	**Post-run**	**p-value**
Kinematic parameters			
DFL (°)	7.0 (2.2)	7.2 (2.7)	0.619
DFR (°)	8.3 (1.9)	7.8 (2.5)	0.279
ICL (°)	1.9 (3.0)	2.2 (3.4)	0.425
ICR (°)	2.7 (2.7)	2.1 (4.1)	0.235
Spatiotemporal parameters			
WS (m/s)	1.24 (0.16)	1.29 (0.18)	**0.003**
Cadence (steps/min)	114 (10)	117 (10)	**0.002**
SLL (m)	0.65 (0.05)	0.66 (0.05)	**0.033**
SLR (m)	0.65 (0.05)	0.66 (0.06)	0.114

Abbreviations: DFL: peak dorsiflexion in swing of left limb; DFR: peak dorsiflexion in swing of right limb; ICL: ankle angle at initial contact of left limb; ICR: ankle angle at initial contact of right limb; SLL: step length of left limb; SLR: step length of right limb; WS: walking speed.

Another objective was explore any gait pattern differences between the two groups. As Table 5.5 shows, there was a statistically significant difference between groups in the pre and post exercise change in peak DF in swing. The MS group was characterised by a mean decrease of -1.5° compared to healthy controls with an increase of 0.2° [t(28)=-2.523, p=0.18, 95% CI (-3.15 -0.32)]. Interestingly, there were no statistically significant differences in any of the kinematic (i.e. notably the peak DF in swing) and spatiotemporal parameters between the MS and the healthy control group before the exercise task.

Table 5.5 Gait kinematic and spatiotemporal differences (post-pre exercise) between pwMS and healthy individuals.

	Difference (post-pre run)		
	MS (n=15)	**HC (n=15)**	**p-value**
Kinematic parameters			
DFMA (°)	-1.5 (2.2)	0.2 (1.5)	**0.018**
DFLA (°)	-0.1 (1.5)	-	0.570
ICMA (°)	-0.2 (2.1)	0.3 (1.6)	0.380
ICLA (°)	0.8 (1.8)	-	0.770
Spatiotemporal parameters			
WS (m/s)	0.02 (0.08)	0.05 (0.05)	0.181
Cadence (steps/min)	1 (3)	3 (2.8)	**0.043**
SLMA (m)	0.01 (0.04)	0.01 (0.02)	0.949
SLLA (m)	-0.001 (0.03)	-	0.315

Abbreviations: DFMA: peak dorsiflexion in swing of most affected limb; DFLA: peak dorsiflexion in swing of least affected limb; GPSMA: gait profile score of most affected limb; GPSLA: gait profile score of least affected limb; ICMA: ankle angle at initial contact of most affected limb; ICLA: ankle angle at initial contact of least affected limb; SLMA: step length of most affected limb; SLLA: step length of least affected limb; WS: walking speed.

5.4.4 Individual analysis of self-reported and objective measures of fatigability for the MS group

The participants of the MS group reported, through a foot drop audit, whether or not they experience transient foot drop and during which activities this phenomenon would occur.

It can be seen from the individual data in Table 5.6 that eight out of the fifteen pwMS reported experiencing transient foot drop, most commonly during running or walking and one participant during cycling. However, objectively, in the second visit a decrease in peak DF in swing exceeding the $MDC_{95\%}$ value (2.4°) was only recorded in three out of the fifteen pwMS while in the third visit this was recorded in six out of fifteen and with one participant exhibiting a deterioration of DF in swing exactly of 2.4°. What is striking about this individual dataset in this table is that four pwMS who had reported that they do not experience transient foot drop exhibited a deterioration in their ankle kinematics with a decrease in peak DF in swing greater than the $MDC_{95\%}$ either in one of the two or in both visits. Moreover, participants (five out of the eight) who had reported experiencing transient foot drop while exercising did not exhibit a deterioration in gait after the 20-minute treadmill run at a self-selected speed in the third visit.

Table 5.6 Self-reported and objective measures of foot drop in the two visits for the MS group.

Participants	**Reported foot drop**	**Activities**	**ΔDF° (post-pre run) 2nd visit**	**Average running speed (m/s) (2nd visit)**	**Peak RPE (2nd visit)**	**ΔDF° (post-pre run) 3rd visit**	**Average running speed (m/s) (3rd visit)**	**Peak RPE (3rd visit)**	**EDSS**
PP1	No	-	-1.6	1.55	15	-0.5	1.63	16	1.5
PP2	Yes	Running	-1.7	1.86	13	3.1	1.91	17	2.5
PP4	No	-	3.0	2.8	13	0.6	3.0	12	1.5
PP5	Yes	Running/cycling	-2.2	2.75	13	**-3.5**	2.41	15	2.0
PP6	Yes	Running	0.2	2.86	17	-0.9	2.58	17	1.0
PP9	No	-	-1.3	1.61	13	**-3.3**	1.86	11	2.0
PP10	No	-	**-3.0**	-	15	**-2.9**	-	13	1.5
PP12	No	-	**-3.1**	2.13	12	-0.8	1.86	12	1.0
PP13	No	-	-1.8	2.63	13	-2.4	2.55	13	2.5

PP15	Yes	NR	2.1	1.3	13	**-4.4**	1.41	12	1.0
PP16	Yes	Running/walking	0.5	2.72	14	1.2	2.63	13	1.5
PP17	No	-	-1.8	2.86	14	**-3.0**	2.8	12	1.0
PP19	Yes	Running	-1.9	2.0	17	0.2	2.19	16	1.5
PP22	Yes	Running/walking	**-7.5**	2.27	13	**-4.7**	2.27	15	2.0
PP23	Yes	Jogging/walking	-2.3	1.69	13	-1.8	1.33	15	2.5

Participants in both visits had similar self-reported responses in RPE, psychological affect and fatigue ratings and average speed over the 20-minute treadmill run, although it the third visit the speed was slightly higher.

5.4.5 Core temperature changes and fatigability

Another objective in this study was to explore the changes in the core temperature during the exercise task. During the second visit, six of the 15 pwMS consented to swallow an ingestible core temperature capsule in order to record the change in core temperature during the 20-minute treadmill run at a self-selected speed. This was done to gain some insight as to whether the change in core temperature would be associated with observations of fatigability throughout the exercise task. Table 5.7 presents core temperature change (ΔT_{core}), absolute end-task temperature (Peak T_{core}), ΔRPE, Peak RPE, feeling scale, ΔVAS Fatigue and average speed during the exercise task for the three who exhibited a deterioration in peak DF in swing above the $MDC_{95\%}$ and the three who did not exhibited such a deterioration. Even though these data is descriptive, these data indicate that the group exceeding the $MDC_{95\%}$ value although selecting to run at a lower speed, had similar core temperature and self-reported RPE, psychological affect (FS) but a possibly slightly higher fatigue response during the exercise task.

Table 5.7 Description of self-reported response characteristics and core temperature during the exercise task of the 2nd visit for the MS group who took the thermistor capsule.

Response characteristics	**DF > $MDC_{95\%}$ (n=3)**	**DF ≤ $MDC_{95\%}$ (n=3)**
ΔDFMA (°)	-2.2	0.5
Average running speed (m/s)	2.19	2.5
Peak T_{core} (°)	38.2	38.4
ΔT_{core} (°)	0.95	1.08
Peak RPE (6-20)	15	17
ΔRPE (6-20)	9	8
Feeling Scale (+5 to -5)	2.8	3.1
ΔVAS Fatigue	3.9	2.6

Abbreviations: ΔDFMA: change in peak DF in swing of most affected leg; T_{core}: core body temperature; RPE: Rating of Perceived Exertion; VAS: Visual Analogue Scale

5.5 Discussion

5.5.1 Pre-post exercise task gait changes in pwMS

The first objective in this study sought to determine the presence of gait deterioration after a 20-minute treadmill run at a self-selected speed in highly active pwMS who were diagnosed with stable, relapsing remitting MS. The most important finding was the statistically significant decrease of peak DF in swing after the exercise task, with six out of the 15 pwMS showing a decrease in DF exceeding the $MDC_{95\%}$ value. This finding is consistent with that of other studies that have investigated kinematic changes after a

6MWT in people whose daily walking ability is affected. The study of McLoughlin et al. (2016) found statistically significant decrease only in DF at initial contact, while the study of van der Linden et al. (2018) demonstrated a statistically significant decrease of both peak DF in swing and at initial contact. However, to our knowledge this is the first study that confirmed the presence of exercise gait deterioration in people with EDSS < 3.0. A 20-minute treadmill running task was selected for the current study as it was thought that the 6MWT would not be sufficient demanding to elicit any gait changes in this highly physically active group.

There is a great amount of literature investigating fatigability evidenced as a decrease in walking speed, the distance walked or in a few studies as a deterioration of gait kinematics over a certain exercise task. The exercise tasks that were most commonly used were the 6MWT or longer distance walking protocols across a wide range of the disability spectrum in MS (EDSS range: 0-6.5) (Sehle et al, 2011; Phan-Ba et al, 2012; Dalgas et al, 2014; Leone et al, 2016; McLoughlin et al, 2016; van der Linden et al, 2018). Thus far, no studies have evaluated fatigability evidenced as foot drop in highly active pwMS with minimal impairments.

Another interesting finding was that we did not detect any statistically significant differences in spatiotemporal parameters (i.e. walking speed, cadence and step length) pre and post exercise task. There are conflicting findings in the literature reporting in spatiotemporal parameters after fatiguing protocols on pwMS. Some studies are in accordance with our findings and did not observe any changes in walking speed, cadence or step length after a 6MWT (Feys et al, 2013; McLoughlin et al, 2016). In the study by Feys et al. (2013) changes in spatiotemporal parameters after a 2MWT and a 6MWT were

investigated in three groups of pwMS with different level of walking ability. It was found that the least impaired group (EDSS range: 1.5-4.0) did not show changes in spatiotemporal parameters pre-post the fatiguing walking task. However, other studies have found either decrease in cadence (van der Linden et al, 2018) or increase in step length and step width after a fatiguing task (Sehle et al, 2011). These conflicting results among studies is likely to be associated with the different fatiguing protocols (i.e. 6MWT, walking until exhaustion, 20-minute treadmill run) and with the level of disability of the MS population that was included in those studies, with increased gait deterioration at higher EDSS levels.

5.5.2 Gait differences between MS and healthy control group

Another objective of the present study was to examine potential differences in (pre-exercise task) gait kinematics and spatiotemporal parameters between the MS and a healthy control group. We found no statistically significant differences in any kinematic or spatiotemporal parameters of gait prior to the exercise task between the MS and the healthy control group.

Previous research studies comparing the gait characteristics of healthy individuals and pwMS with minimal neurological disability (EDSS < 3.5) have reported ambiguous findings. Consistent with our findings of no changes (before the exercise task) in the spatiotemporal parameters was a single study that reported no significant differences in walking speed, step length and double support time between pwMS (EDSS range: 0-1.5) and healthy individuals (Nogueira et al, 2013). However, in contrast to our findings, most other studies demonstrated that pwMS with EDSS between 0 to 2.5 walked slower, with

associated decreased cadence, stride and step length and prolonged double support time (Benedetti et al, 1999; Martin et al, 2006; Kalron et al, 2011; Sosnoff et al, 2012). A recent meta-analysis summarizing the gait deficits in pwMS (EDSS range: 1-4.8) compared to healthy individuals found large effects of MS on walking speed [Standardised Mean Differences (SMD = 1.12)], stride (SMD = 1.27), step length (SMD = 1.15), double support (SMD = 0.85) and swing phase duration (SMD = 1.23) with all effect sizes increased when participants were asked to walk at faster speeds (Comber et al, 2017).

Our findings also did not demonstrate any statistically significant differences in the pre-exercise kinematic parameters between the MS and the healthy control group. Interestingly, these findings are not supported by other studies. A number of studies have investigated gait kinematic characteristics of pwMS that have minimal neurological deficits (EDSS < 2.5) and it was observed that compared to healthy individuals pwMS had a significant decrease in dorsiflexion at initial contact, at toe off and DF during the swing phase of gait (Benedetti et al, 1999; Martin et al, 2006; Nogueira et al, 2013). These conflicting findings between our results and the rest of the studies could be associated with the high physical activity status of our MS group compared to the MS population from the other studies. Even though pwMS in the other studies also had minimal neurological deficits (EDSS < 3) and were without any mobility impairments, our study population consisted of active pwMS that were exercising at least two or three times in a week. The association between physical activity level and walking performance in pwMS has been documented (Snook et al, 2009; Cavanaugh et al, 2011). For example, Cavanaugh et al. (2011) reported that pwMS without any mobility limitations (EDSS < 4.0) accumulated an average of 8860 steps daily which is consistent with levels recorded

in the general adult population. The pwMS in our study however seem to be far more physically active with an average of over 11000 steps a day. One may speculate that a higher physical activity level and associated increased fitness levels may have led to the lack of even mild deterioration in gait kinematics in our participants (pre-exercise).

An important finding was the significant decrease of peak DF in swing in pwMS compared to our healthy control group. This could be a sign of exercise-induced fatigability in the MS population that does not occur in the healthy individuals. It is likely that the cause of exercise-induced fatigability in MS is multifactorial. It has been suggested that fatiguing protocols in pwMS induce impaired central motor activation (Andreasen et al, 2009), while other studies reporting increased central activation during fatiguing exercises which could possibly reflect a compensation mechanism (Thickbroom et al, 2008). Thus, it could be that the observed deterioration of the gait parameters after fatiguing protocols is a reflection of the failure of these compensatory mechanisms. Further, another possible explanation for this significant decrease in peak DF in swing could be the Uhthoff's phenomenon that is well known to trigger a number of symptoms in MS (Frohman et al, 2013). The demyelination interferes with the transmission of the nerve impulses along the axons and even though there are compensatory mechanisms to restore the nerve impulse conduction in demyelinated axons, these axons are more susceptible to changes such as rise in core temperature (due to exercise) (Lassmann et al, 2012).

5.5.3 Individual analysis of self-reported and objective measures of the second and third visit for the MS group

The results of this study indicated that the MS group in third visit had a significant decrease in peak DF in swing (≈1.5°) after completing the exercise task. What is interesting is that by observing the individual performance of each participant, in the second visit only three out of 15 exceeded the $MDC_{95\%}$ value, while in the third visit six out of fifteen showed a worsening exceeding the $MDC_{95\%}$. What is more interesting is that in only two participants a decrease in peak DF in swing exceeding the $MDC_{95\%}$ was observed in both the second and third visit. Several factors could explain this observation. Firstly, since participants had three visits, with the first visit consisting of habituation on running/jogging on the treadmill and the following two consisting of completing the same 20-min treadmill run at a self-selected speed, there could be a certain amount of training effect which resulted in a slightly higher running speed on the exercise task in the third visit with possible associated worsening of the walking afterwards. An additional aspect that should be taken into account is the increasing motivation of the participants to increase their running speed since they were more accustomed and comfortable to put more effort in the exercise task than the previous visit, which as explained above could have led to more participants experiencing fatigability. Another possible explanation for the difference between the two visits is the within-day and day-to-day variability of gait parameters in pwMS, although evidence for this is conflicting. For example, Albrecht et al. (2001) reported marked day-to-day variability in individual maximum walking time and distance of unaided walking in pwMS with EDSS range from 4.0 to 5.5. A study by Feys et al. (2014) examined the within-day variability of gait parameters during short and long distance walking tests in a wider range of disability level among pwMS (EDSS range: 0-6.5) demonstrated that there was greater within-day variability over shorter distance

walking tests and the variability increased proportionally with the disability level. Especially for the least impaired group the within-day variation of walking speed was far below 20% (Feys et al, 2014). However, Morris et al. (2002) reported that from morning to afternoon there was very little change in any of the spatiotemporal parameters in either MS or healthy populations, even though self-reported fatigue was increased in the afternoon in pwMS. Similarly, Crenshaw et al. (2006) reported that there were no significant changes of kinematic and spatiotemporal parameters variability in a group of pwMS from morning to afternoon after a 15min walk, even though self-reported fatigue was increased in the afternoon. The majority of these studies have investigated the within-day variability, while the variability in detection of fatigability is between-day. Further, these studies, including our reliability study in gait kinematics and spatiotemporal parameters described in Chapter 4, have reported on the variability in gait characteristics and not in the exercise-induced change of gait characteristics.

5.5.4 Perceived versus objectively recorded fatigability

Another interesting finding in this group of pwMS was that in the third visit six out of the 15 exhibited post-exercise reduction in peak DF in swing exceeding the $MDC_{95\%}$. Three out of the six had reported that were not experiencing transient foot drop while exercising, even though it was objectively documented with the reduction in peak DF in swing post-exercise. This rather contradictory result might be due to the fact that these are highly physically active people and subtle changes could not be easily perceived, whilst with 3D gait analysis, as shown to be the 'gold' standard, even subtle differences can be recorded. On the other hand, five out of the fifteen had reported in the audit that they experience

'transient' foot drop while exercising, but this was not objectively documented. It is difficult to explain this result, but it might be related to the physical performance on the exercise task where the 20-min treadmill run might not have been long enough to elicit the phenomenon.

5.5.5 Core temperature and self-response vs objectively measured characteristics

A sub-analysis from the six participants that accepted to ingest the thermistor capsule for recording of the core temperature while performing the exercise task showed that the core temperature change was an increase of around ≈1.0°C. Similar to the thermal responses observed in our sub-group, studies have shown core temperature changes of ≈0.9±0.4°C after endurance exercise (Skjerbaek et al, 2012; Sandroff et al, 2016). However, these studies did not investigate changes in gait kinematics or spatiotemporal parameters. In this sub-group, three pwMS showed changes exceeding $MDC_{95\%}$ and three of them did exhibit this worsening in peak DF in swing. Although because of the small sample size no inferential statistics could be performed, both groups seemed to demonstrate similar core temperature and self-reported RPE and fatigue rating responses. This might indicate that heat sensitivity per se was not a factor influencing the reduction in peak DF in swing. However, future studies should look into heat sensitive pwMS and exercise-induced gait changes and confirm or refute these preliminary findings.

5.5.6 Limitations and future suggestions

Some methodological considerations need to be addressed. Firstly, this experimental study had a small sample size and power calculation was not performed in order to detect

changes in gait characteristics. However, it is aimed to act as a pilot study in order to be able to calculate the sample size that would be needed to conduct a longitudinal multi-centre trial. Further, another possible limitation that we could not control was that some of the participants could not attend the second and third session at the same time of the day. As mentioned earlier, the time of the day could have influenced the performance on the exercise task depended on the fatigue level of the individuals and subsequently have an effect on the walking ability prior and after the exercise task. Further, as only six out of fifteen agreed to ingest the thermistor capsule, our preliminary finding that a change in core temperature is not associated with the presence of fatigability measured as a reduced peak DF in swing needs to be confirmed in future investigations. Such investigations may also record the change in core temperature in healthy controls.

Since this was a pilot study we used the data obtained to inform sample size calculations for future definitive studies based on the same primary outcome. Based on a calculation on g*power (Faul et al, 2017) with an effect size of 0.44 for the peak DF in swing, with 80% power to detect statistical significance ($p < 0.05$) between pre-post measurements, a number of 33 participants will be needed.

5.6 Conclusion

The purpose of the current study was to investigate the presence and prevalence of exercise-induced fatigability, evidenced as deterioration in gait kinematics after an exercise-running task in stable, relapsing, highly physically active pwMS who presented with minimal or very mild neurological disability (EDSS < 2.5). A secondary aim was to compare before and after an exercise task the gait of these highly active pwMS with a

healthy control group who were also highly active. It was found that there is a significant decrease of peak DF in swing of the most affected limb in the MS group post-exercise. Inspecting the individual data of the MS group showed that six out of the 15 in the MS group showed reduced peak DF in swing exceeding the $MDC_{95\%}$ value. However, unlike previous studies we did not observe any significant differences in the gait kinematics and spatiotemporal parameters in our participants with high levels of habitual physical activity compared with our healthy controls. Although highly speculative, this finding could indicate that high levels of physical activity protect against gait abnormalities. Future studies should consider investigating potential changes in other joints and planes, apart from sagittal ankle kinematics, which could possibly reveal different compensatory mechanisms for foot drop. As the focus of the present study was to examine fatigability evidenced as foot drop, we reported only sagittal ankle kinematics.

Moreover, future studies should consider the use of a more standardized fatiguing task, so that the effort of all the participants can be relativised. Thus far, no standardised protocol is available to induce fatigability in pwMS, with lack of information on psychometric properties and comparison of different protocols. A standardised protocol should take into account the physical activity and disability level of pwMS and should be relativized according to each individual's peak capacity.

Finally, once a protocol that can reliably detect fatigability has been developed, this could potentially monitor longitudinally the worsening of gait kinematics after exercise-induced fatigability, through 3D gait analysis, since subtle changes in gait could possibly be used to detect changes in disease progression in this sub-patient population.

Chapter 6. The direct orthotic effect of FES on gait kinematics and walking speed in people with MS under dual-tasking and fatiguing walking conditions

6.1 Purpose of the chapter

The purpose of the present chapter is to investigate the direct orthotic effect of FES under a variety of walking conditions in people with MS who were prescribed FES to treat their foot drop. It was hypothesized that these walking conditions would simulate real life situations in which people have to focus on more than one task at the same time with and without exercise induced fatigability. And thus, the perceived benefits of the use of FES that pwMS have previously reported.

6.2 Introduction

As discussed previously, foot drop, i.e. the lack of dorsiflexion during the swing phase of gait is a common symptom in pwMS (Mount & Dacko et al, 2006). A common way to treat foot drop is through FES, where pre-tibial muscles are stimulated in order to produce ankle dorsiflexion. It has been reported that FES facilitates a more normal gait pattern and improves perceived quality of life (Stein et al, 2010). Many studies have investigated the direct orthotic effect of FES (FES off vs FES on) and reported an increased walking speed and a decrease in energy consumption cost (Taylor et al, 1999; Stein et al, 2010; Scott et al, 2013; Street et al, 2015). On the other hand, studies investigating the long-term (therapeutic) effect of FES to treat foot drop in pwMS, did not observe improvement over

time (Taylor et al, 1999; van der Linden et al, 2014a; Street et al, 2015), which might possibly be attributed to the progressive nature of the disease.

Fatigue is an important factor that negatively affects the everyday life of pwMS. According to the taxonomy by Kluger et al (2013) there is a difference between perceptions of fatigue, which refers to the weariness and self-reported exhaustion, and fatigability, which is the exercise induced decline in the ability of the muscles to produce force or power regardless of whether the task can be sustained. To the best of our knowledge, only one study investigated how the use of FES affects fatigability-induced gait changes. The study by Barr et al. (2017) observed that an orthotic effect was not evident in the first use of FES, but after eight weeks of FES use there were positive effects on gait with the use of FES compared to without. However, in that particular study, the use of FES alone was not sufficient to overcome the deficits in gait related to fatigue after a 6MWT task.

One of the benefits of the use of FES reported by pwMS themselves is that it reduces the mental effort of walking, as less concentration is needed on the walking task (Bulley et al, 2015; (Miller) Renfrew et al, 2018). It has been suggested that postural control and cognition compete for a common pool of attentional resources and if one task is becoming more challenging, the available resources reach their limit and performance in one or both tasks will deteriorate (Stins & Beek, 2012). The execution of a motor task simultaneous with a cognitive task is termed dual-tasking and the cognitive-motor interference can be quantified by the dual-task cost (DTC), which is the percentage change in performance from single to dual task (Yogev-Selingmann et al, 2012).

Several studies have indicated that while performing a cognitive task together with a walking task, pwMS exhibited deterioration in gait evidenced as reduced walking speed and associated reduced step length and increased double support time (Hamilton et al, 2009; Wajda et al, 2013). The detrimental effect of the cognitive task on walking performance was observed to be more severe in pwMS with higher levels of disability (Sosnoff et al, 2011a), but was also evident in people with clinically isolated syndrome (Kalron et al, 2010).

Although many studies have investigated dual-tasking in pwMS, the influence of FES on the dual-task cost and the orthotic effect of FES have not been explored. Miller et al. (2016) showed that, unlike patients with walking speeds lower than 0.8m/s, for those with higher walking speeds the use of FES does not induce an increase in walking speed during laboratory timed walking tests. However, subjectively, people with walking speeds faster than 0.8m/s, i.e. near normal walking speed, still report to benefit from the use of FES. In the current study, it was hypothesized that the benefit perceived in those people may be due to benefit of using FES when experiencing fatigability and/or when performing another task, i.e. when dual tasking.

The aim of the present study was to investigate the direct orthotic effect of FES under conditions which may be more ecologically valid than a single task, short duration laboratory walking test conventionally used to investigate the effects of FES in gait. The current study compared three different conditions: 1) a single task walking condition without fatigability (task A), 2) a dual-task walking condition without fatigability (task B) and 3) a dual-task walking condition with fatigability (task C).

• The first objective was to explore the direct orthotic effect of FES in these three conditions on peak DF in swing and walking speed. We hypothesized that the direct orthotic effect of FES would be higher in task B compared to task A and also would be higher in task C compared to task A.

• The second objective was to quantify the DTC of walking speed of task B of the MS group and to compare it with the DTC of a healthy age-matched control group. Based on other studies (i.e. Hamilton et al, 2009) it was hypothesized that the DTC of walking speed for the MS group would be at least 10%.

• The final objective was to investigate the effect of fatigability (without the assistance of FES) measured as a deterioration in gait characteristics in task C (dual-task and exercise-induced fatigability) compared to task B (dual-task only) in both groups. It was also hypothesized that gait deterioration would be greater in the MS compared to the healthy control group after inducing fatigability and performing a cognitive task.

6.3 Methods

6.3.1 Design

This was a non-randomised experimental study to investigate the orthotic effect of FES under a variety of walking conditions in pwMS who experience foot drop. Data collection for this study started in August 2018 and is still ongoing until the sample size derived from the *a priori* power calculation is met.

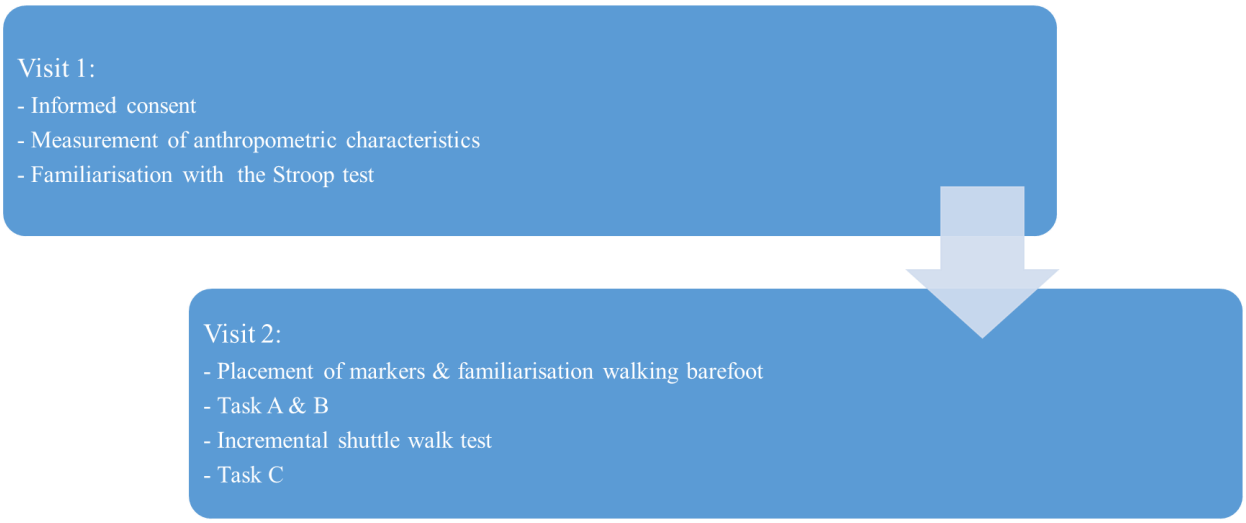


Figure 6.1 Flow diagram of study design for Chapter 6.

6.3.2 Participants

Eligibility criteria required pwMS to have received clinically definite diagnosis of MS according to the revised McDonald criteria (Polman et al, 2011), be aged 18 years and over and to be characterized with moderate to severe disability according to EDSS (EDSS ≤ 6.0). They should also be able to walk at least 100 meters with or without walking aids. Further, participants should be community walkers who experience foot drop diagnosed by a physiotherapist and were using FES to treat foot drop. Participants were excluded from the study if they had a clinically diagnosed relapse within the last month and had any musculoskeletal impairments and cardiopulmonary disorders that could affect their walking ability.

In order to be eligible in the healthy control group, participants were aged 18 and over and were also 'free' from the diagnosis of any neurological disease or any other condition/injury that would affect their walking ability.

6.3.3 Ethical opinion

Favourable ethical opinion was obtained from the NHS Research Ethics Committee (West Midlands – Edgbaston Research Ethics Committee, REC Reference number: 18/WM/0062) (Appendix 15) and Queen Margaret University Ethics Committee. Moreover, NHS Lothian Research and Development approval was obtained prior to commencing with the study (Appendix 16). All procedures were conducted in accordance with the Declaration of Helsinki with regards to human participation. Participants could withdraw from the study at any point without giving any reason.

6.3.4 Recruitment process

People with MS who use FES to treat foot drop were recruited from the MS Therapy Centre Lothian, the Anne Rowling Clinic and the Physiotherapy Department at Astley Ainsley Hospital, all based in Edinburgh. The specialist neurologists, physiotherapists and nurses from these centres identified potentially eligible participants. If people were interested to learn more about the study, a participant information pack was provided to them including a PIS (Appendix 17) and the self-report questionnaires [Multiple Sclerosis Neuropsychological Screening Questionnaire (MSNQ), foot drop questionnaire, FSMC & MSWS-12]. In the pack, we also included a stamped addressed envelope addressed to the PI (PhD student) and a form stating their agreement to be contacted by the PI. The potential participants were given at least two weeks to consider the information and contact the PI, their GP or other independent parties to decide whether they would take part. On receipt of the form and the screening questionnaires, the PI contacted those

eligible to take part in order to discuss the study protocol, answer any questions and where appropriate arrange the first study visit. On arrival at QMU motion analysis laboratory for the first visit, participants were asked to sign an informed consent form (Appendix 18). It was clearly stated that they could withdraw from the study at any point with no obligation to give any reason and without prejudice to their future care.

The healthy control group of participants was age and gender-matched with the MS group and was recruited through word of mouth, through the QMU Moderator system and by advertisements in the recruitment centres (Appendix 19).

6.3.5 Study protocol

The study protocol required the participants in the MS group to visit the QMU motion analysis laboratory on two occasions.

6.3.5.1 Visit one

On arrival at the motion analysis laboratory, the study protocol was explained to the participants in detail and participants were given the opportunity to ask any questions. The next step was to sign the informed consent prior to commencing with any procedures. The purpose of this visit was to record anthropometric characteristics required for the processing of 3D gait analysis data, such as height, body mass, leg length, knee and ankle width and tibial torsion. If participants reported that they were experiencing difficulties with their use of FES, the PI would advise them to contact the physiotherapists at Astley Ainsley at the FES clinic for a follow-up appointment

Figure 6.2 Stroop test

regarding their FES device. Finally, the Stroop test (Franzen et al, 1987), which they would have to perform on their second visit was explained. The Stroop test is an attention-demanding task that 'assesses the ability to inhibit cognitive interference which occurs when the processing of a stimulus affects the simultaneous processing of another attribute of the same stimulus' (Stroop, 1935). In this test, participants were shown a word ('blue', 'yellow', 'red' or 'green') written in a different colour (blue, yellow, red or green) and were instructed to state the colour of the text and ignore the meaning of the word itself (Figure 6.2). The font of the letters was large enough for the participants to read from a distance of seven meters and a new word was displayed every second. In this visit, participants performed a few practice trials to become accustomed with the protocol and to understand the task that they would be asked to perform in the next visit. This first visit lasted approximately one hour.

6.3.5.2 Visit two

The second visit to the motion analysis laboratory lasted approximately two hours. The purpose of this visit was to perform the walking (single and dual-task trials) and physically demanding procedures expected to induce fatigability to assess the direct orthotic effect of FES. On arrival, reflective markers were placed on the lower limbs and pelvis of the participants according to the Helen Hayes' marker system (Kadaba et al, 1990). These small (9mm diameter) markers were attached using skin friendly double-sided tape, in order to allow the 8-camera 3D motion capture system (Vicon Motion Systems, Oxford, UK) to record the kinematic and spatiotemporal parameters of the lower limbs and pelvis (see details in section 4.3.5). The walking trials had to be performed barefoot and to ensure

that the foot switch from the FES device would still work we had to attach it with adhesive tape to the sole of the foot. Participants were given a few walking trials to familiarise themselves walking barefoot with the foot switch attached on the sole of the foot. Afterwards, participants were asked to complete a total of 24 short (approximately 7m) walking trials on three different conditions (see Table 6.1). The walking trials of the first and second condition (single-task vs dual-task) with FES on and off were counter-balanced, so that a potential order effect was avoided. In order to do so, each trial was allocated a specific number and the sequence of the walking trials was decided by computer-generated randomization. Participants were allowed to rest between walks before (but not after) performing the fatiguing task.

Table 6.1 The three different conditions in the study protocol. The order of trials with FES on and off were randomized.

Condition	**Walking trials**
Task A: single-task walking	• Walking with FES on (4 trials) • Walking with FES off (4 trials)
Task B: dual-task walking with the Stroop test	• Walking with FES on (4 trials) • Walking with FES off (4 trials)
Task C: dual-task walking with the Stroop test after induced fatigability	• Walking with FES on (4 trials) • Walking with FES off (4 trials)

*The healthy control group performed all the conditions except those walking trials with FES on.

After the completion of the walking trials of the first two conditions, the fatiguing task was performed. The fatiguing test that was chosen was the shuttle walk/run (SRT) test that was developed for children with CP at GMFCS level II (Verschuren et al, 2006). The initial plan was to use the Incremental Shuttle Walk Test (Singh et al, 1992) which is an incremental shuttle exercise test originally developed for patients with chronic obstructive pulmonary disease. However, the starting speed of that test was 1.80km/h with a 0.60km/h increase every minute, which after pilot testing was evident that the MS group would not be able to perform for more than a few 'shuttles'. Hence, it was decided that the best method to induce fatigability in this population was the SRT at GMFCS level II since the initial speed is at 2km/h and the speed is increased by 0.25km/h every minute. The SRT-II, comprised of 23 levels, requires participants to walk between two cones (10m distance) in time between two bleeps sounding from a CD player. Every minute, the time between the two bleeps is shorter hence participants were required to walk at a faster speed. The test was terminated if participants indicated that they could not continue with the test at the speed required. Furthermore, for safety reasons they were allowed to use their walking sticks and FES to complete the task. Directly after the termination of the SRT-II, participants performed the last eight walking trials (four with FES on and four with FES off) while performing the Stroop test.

The participants in the healthy control group were asked to visit QMU motion analysis laboratory only for one session, in which all the procedures described above took place. At first, the purpose of the study was explained after which informed consent was taken.

Anthropometric measurements were recorded and participants were given the opportunity to practice with the Stroop test. This group performed the single and dual-task conditions in a counterbalanced order, as in the MS group, but of course without the walking trials with FES, hence performed a total of 12 walking trials.

6.3.6 Outcome measures

6.3.6.1 Gait analysis

As described in the previous chapters of this thesis, 3D gait analysis is the 'gold' standard in terms of quantitative gait analysis and it has been widely used for quantifying gait changes in the MS population. In the present study, gait analysis was performed while participants performed 24 (MS group) or 12 (healthy control group) barefoot walking trials over a distance of about 7 meters. The most affected leg was defined as the leg for which they used FES to treat their foot drop. For the healthy control group we selected the left leg for comparison with the MS group, assuming there were no differences between left and right leg for the healthy group. For further details on marker placement and data processing please refer to Chapter 4 in sections 4.3.5 and 4.3.6.

6.3.6.2 Fatigue Scale for Motor and Cognitive Function

The FSMC is a self-reported questionnaire consisting of 20 items assessing motor and cognitive fatigue (Penner et al, 2009) (Appendix 13). Participants were asked to complete the FSMC once at the beginning of the study. For further details regarding the questionnaire, such as scoring and psychometric properties, please refer to Chapter 5, section 5.3.6.5.1.

6.3.6.3 Multiple Sclerosis Walking Scale

The MSWS-12 consists of 12 items assessing the limitations to walking due to MS the past two weeks (Appendix 20). Each item is scored from 1 (not at all) to 5 (extremely) with the total score ranging from 12 to 60 and can be transformed to a percentage, with higher scores indicating greater impact on walking (Hobart et al, 2003). The psychometric properties of the MSWS-12 have been extensively examined. In Chapter 3, it was found that the MSWS-12 has a strong level of evidence for internal consistency, test-retest reliability and construct validity (see Andreopoulou et al, 2018).

6.3.6.4 Multiple Sclerosis Neuropsychological Screening Questionnaire

The MSNQ is a 15-item self-reported questionnaire with scores ranging from 0 (never, does not occur) to 4 (very often, very disruptive) for each item and with a value of 27 and above indicating cognitive impairment (Appendix 21). The MSNQ was designed as a screening tool that is easy and quick to administer by a nonprofessional and it is based on self-report observations over the past three months (Benedict et al, 2003). It has been found to be a reliable and valid tool for screening in the MS population (Benedict et al, 2003; Nauta et al, 2018) and it has been translated in many languages (Sonder et al, 2012; Sejbæk et al, 2018).

6.3.6.5 Hospital Anxiety and Depression Scale

The Hospital Anxiety and Depression Scale (HADS) is a 14-item self-report measure divided in two subscales (depression and anxiety) (Appendix 22). The scoring for each

subscale can range from 0 to 21 and a score of 8 or above on either of the subscales can indicate possible anxiety or depression (Zigmond & Snaith, 1983). It is an easy and quick scale to administer and participants completed it once during the first visit. The HADS has been found to be a valid measure to detect depression and anxiety in the MS population and has been shown to have high sensitivity and specificity for both subscales (Watson et al, 2014; Patten et al, 2015).

6.3.7 Statistical analysis

6.3.7.1 Sample size calculation

In the present study, sample size calculations estimated that 15 participants were needed, using a single group repeated measures ANOVA, to detect a statistical significant ($p < 0.05$) difference between any of the three conditions with effect size of 0.33 and 80% power (overall F-test of difference between any of the three conditions). This was based on the primary outcome that was the direct orthotic effect of FES on ankle kinematics.

6.3.7.2 Data analysis

The initial plan of the descriptive and statistical analysis was as follows. Parametric assumptions of the gait data and walking speed would be checked. A Shapiro-Wilk test would be carried out, to explore the notion that the data are normally distributed. This would be also supported by visual inspection of the q-q plot and boxplot. For the primary question, to explore the orthotic effect of FES in three different conditions a one-way repeated measures ANOVA would be used. Partial eta^2 would also be calculated to determine the effect size.

For the rest of the research questions, since only the walking conditions without FES would be used, paired or independent t-test or the non-parametric equivalent would be performed.

Further, the median and the interquartile range (IQR) was reported for all of the outcomes.

For our primary aim, to explore the direct orthotic effect of FES in the three conditions we calculated the effect size (Cohen's d) on ankle kinematics and walking speed between Task A and Task B and also between Task A and Task C.

For our second aim which was to compare the DTC of walking speed between the MS and the healthy control group we report the mean and standard deviation. The DTC of walking speed for both groups was calculated using the following formula:

$$DTCws = \frac{WS\ single\ task - WS\ dual\ task}{WS\ single\ task} \times 100$$

For the final aim to investigate fatigability-induced changes in both groups, the median and IQR are reported for the conditions of dual tasking before and after the fatiguing (incremental shuttle) task (i.e. Task C vs Task B).

6.4 Results

6.4.1 Participants

A total number of eight pwMS who were FES users and met the inclusion criteria were recruited for this study up to April 2019. One participant had to withdraw from the study after attending the first visit, due to an accident that affected her walking ability. Another

participant performed only three walking trials for each of the conditions instead of four, because of her walking difficulties. Further, seven healthy controls volunteered to take part, but due to issues with the recording of gait kinematics of two, only five were used for comparison with the MS group. Thus, seven pwMS and five healthy age-matched controls were used for analysis in the present study. All participants completed the study protocol without experiencing any adverse effects. As this sample size was below that of an appropriately powered study, no inferential statistical analysis was performed.

Table 6.2 Demographic characteristics of MS and healthy groups presented as means (SD).

	MS group (n=7)	**Control group (n=5)**
Female/Male, n	5/2	4/1
Age, years	54.1 (10.4)	55.0 (11.6)
EDSS range	4-6	-
MSWS-12 (12-60)	46.6 (11.9)	-
MSWS-12 (%)	72.0 (24.9)	-
$FSMC_{tot}$ (20-100)	60.5 (4.6)	26.3 (8.5)
$FSMC_{cognitive}$ (10-50)	26.7 (7.2)	13.0 (4.8)
$FSMC_{physical}$ (10-50)	33.8 (5.0)	16.3 (8.1)
MSNQ (0-60)	20.1 (11.7)	14.7 (7.6)
$HADS_{depression}$ (0-21)	5.8 (2.6)	3.5 (2.9)
$HADS_{anxiety}$ (0-21)	5.8 (3.9)	5.5 (3.7)

Abbreviations: FSMC: Fatigue Scale for Motor and Cognitive Function; HADS: Hospital Anxiety and Depression Scale; MSNQ: Multiple Sclerosis Neuropsychological Screening Questionnaire; MWSW-12: Multiple Sclerosis Walking Scale; PP: Primary Progressive; RR: Relapsing Remitting; SRT: Shuttle Run Test; SP: Secondary Progressive.

The two groups had a similar female to male ratio and average age in years (Table 6.2). All the pwMS were currently FES users, with an EDSS range of 4 to 6. Based on the MSWS-12 scores, the walking ability of most pwMS was moderately affected with an average score of 72% indicating the most severe walking limitations. The MS group was suffering from severe physical fatigue ($FSMC_{physical} \geq 32$), but mild to moderate cognitive fatigue based on the FSMC sub-scales. Further, there were low average scores for cognitive impairment, depression and anxiety based on the MSNQ and HADS (Table 6.2). However, two of the participants in the MS group were exceeding the cut-off value of 27 for the MSNQ. The healthy control group had low scores in all the questionnaires compared to the MS group and completed more shuttles in the SRT to induce fatigability.

6.4.2 Direct orthotic effect of FES

The main objective of the present study was to investigate the direct orthotic effect of FES during three different conditions. The first condition was to walk with FES on and off (Task A), the second condition was to walk with FES on and off while performing the Stroop test (Task B) and the third condition was to walk with FES on and off after completion of a fatiguing task while performing the Stroop test (Task C). Table 6.3 presents the median and IQR of the ankle kinematic and spatiotemporal parameters over the three conditions. It is evident from this table that walking with FES-on resulted in an improvement in most gait parameters, i.e. a direct orthotic effect. Interestingly, in the

conditions A and B, on average, the ankle angle at initial contact was less (more plantar flexed) with the assistance of FES compared to no FES.

The direct orthotic effect of FES was higher in both Tasks B and C compared to Task A (Table 6.4). As shown in Table 6.4, there was a small but positive effect size for the direct orthotic effect for peak DF in swing in both comparisons and for walking speed when comparing Task B and Task A, which indicates that that there was a benefit of FES as hypothesized. However, there was a small negative effect size for both orthotic effects of ankle angle at initial contact and walking speed when comparing Task C to Task A which is the opposite of our hypothesis, reflecting that there was no benefit of FES after the fatiguing task.

Table 6.3 Median (IQR) of gait kinematics and spatiotemporal parameters of the most affected leg in the three conditions (with FES off and FES on the MS group).

	Task A		**Task B**		**Task C**	
	FES off	**FES on**	**FES off**	**FES on**	**FES off**	**FES on**
DFMA (°)	3.1 (10.1)	6.8 (7.2)	1.3 (11.8)	5.9 (6.3)	0.3 (9.4)	6.9 (6.1)
ICMA(°)	1.0 (14.1)	-0.1 (5.1)	1.1 (14.1)	-0.3 (5.2)	0.3 (12.3)	0.5 (7.4)
WS (m/s)	0.8 (0.67)	0.88 (0.6)	0.76 (0.59)	0.8 (0.45)	0.57 (0.41)	0.69 (0.4)
Cadence (steps/min)	99 (65)	104 (51)	98 (67)	95 (46)	89 (48)	95 (48)
SLMA (m)	0.47 (0.1)	0.50 (0.13)	0.43 (0.09)	0.48 (0.16)	0.45 (0.17)	0.45 (0.16)

Abbreviations: DFMA: peak dorsiflexion in swing of most affected leg; ICMA: ankle angle at initial contact of most affected leg; IQR: Interquartile range; SLMA: step length of most affected leg; WS: walking speed.

Table 6.4 Direct orthotic effect of FES in the three conditions [presented as median (IQR) and effect size].

	Task A	Task B	Task C	ES Task A/B	ES Task A/C
Orthotic effect DF (°)	3.7 (4.5)	4.2 (5.2)	4.2 (7.6)	0.12	0.10
Orthotic effect IC (°)	4.6 (11.3)	5.1 (12.4)	4.9 (14.7)	0.04	-0.17
Orthotic effect WS (m/s)	0.02 (0.07)	0.03 (0.13)	0.02 (0.05)	0.26	-0.09

Abbreviations: DF: peak dorsiflexion in swing; ES: effect size; IC: ankle angle at initial contact; IQR: Interquartile range; WS: walking speed.

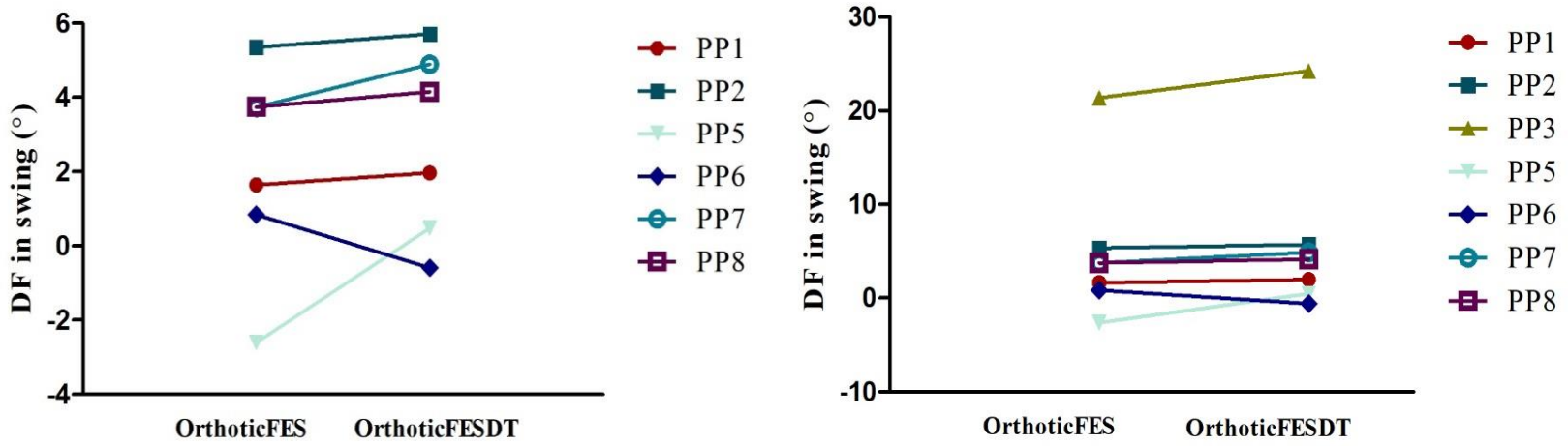


Figure 6.3 The direct orthotic effect on peak DF in swing between Task A (OrthoticFES) and Task B (OrthoticFESDT). Left: all MS participants included; Right: without participant PP3.

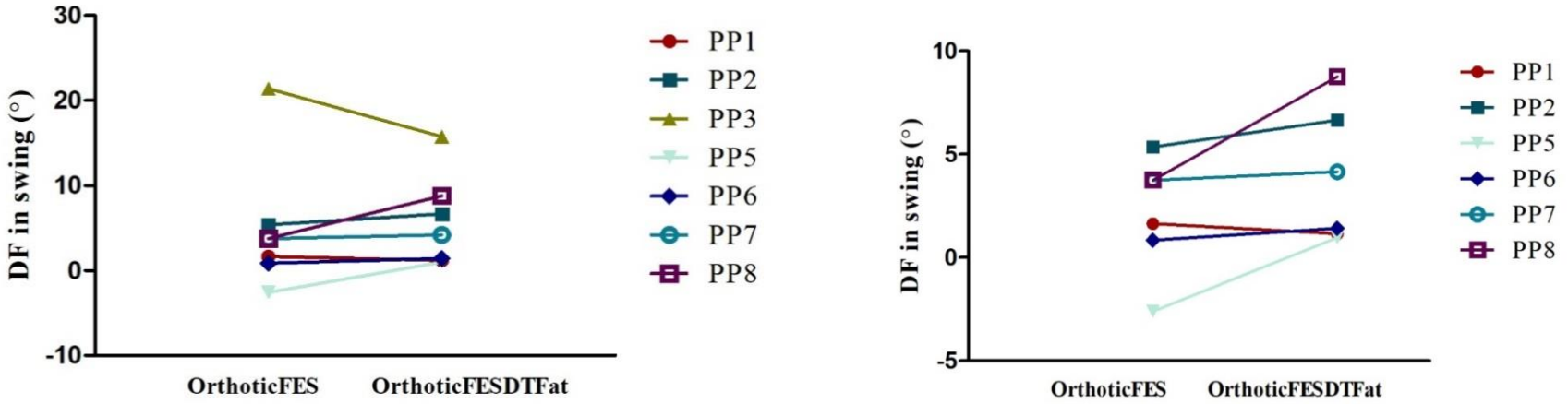


Figure 6.4 The direct orthotic effect on peak DF in swing between Task A (OrthoticFES) and Task C (OrthoticFESDTFat). Left: all MS participants included; Right: without participant PP3.

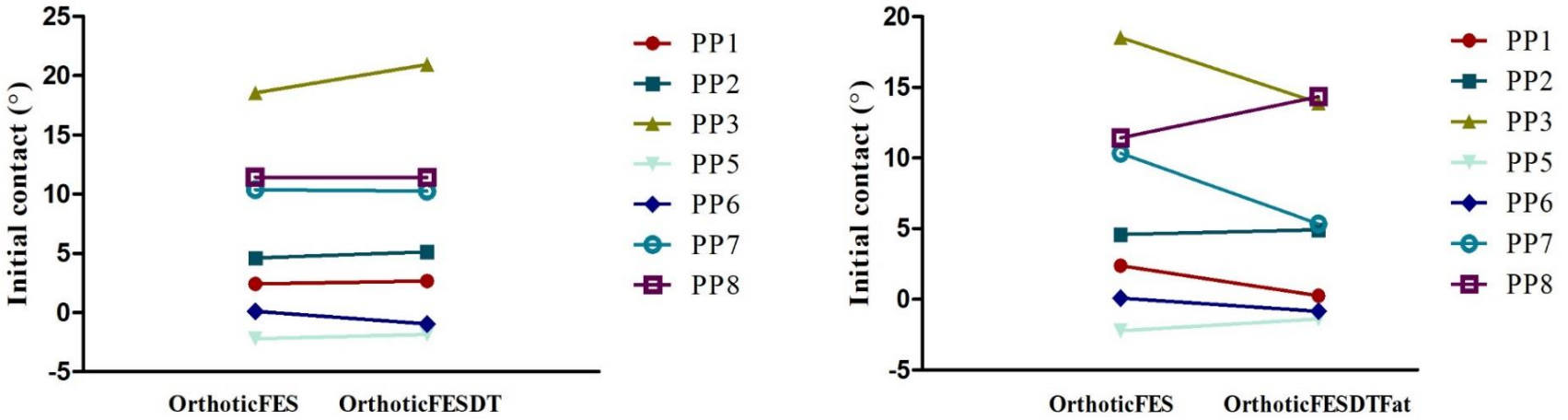


Figure 6.5 The direct orthotic effect on initial contact. Left: between Task A (OrthoticFES) and Task B (OrthoticFESDT); Right: between Task A (OrthoticFES) and Task C (OrthoticFESDTFat).

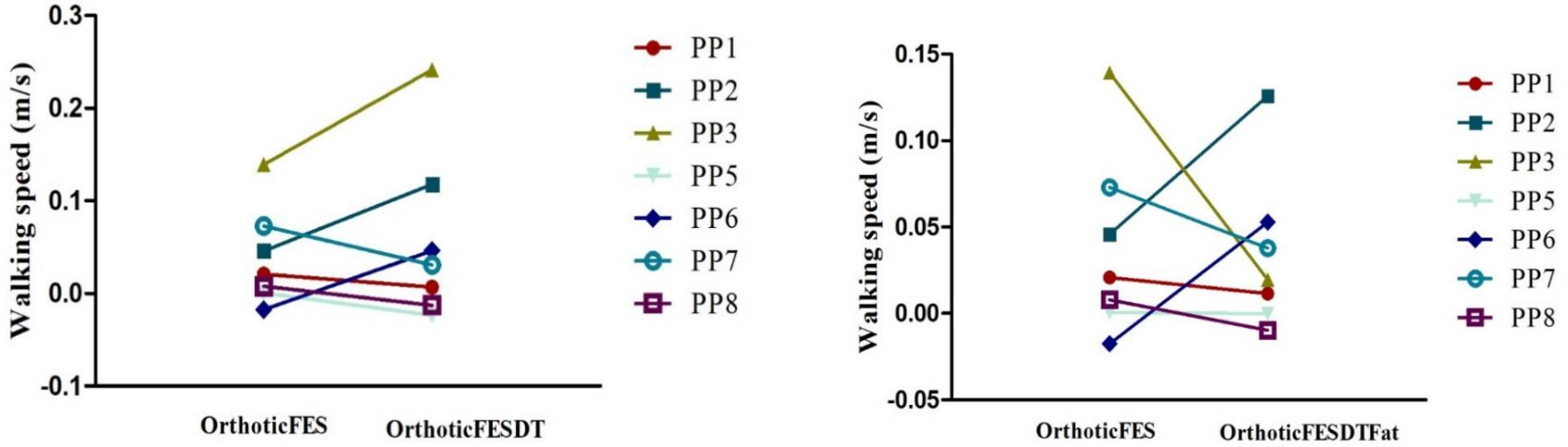


Figure 6.6 The direct orthotic effect on walking speed. Left: between Task A (OrthoticFES) and Task B (OrthoticFESDT); Right: between Task A (OrthoticFES) and Task C (OrthoticFESDTFat).

Figures 6.3 to 6.6 presents the individual comparisons of the direct orthotic effect of FES between Task A/B and Task A/C for peak DF in swing, ankle angle at initial contact and walking speed. For peak DF in swing, six out of seven had a higher direct orthotic effect during the dual-task (Task B) compared to the single-task condition (Task A) whilst five out of seven had a higher direct orthotic effect after the fatiguing condition (Task C) compared to the single task condition (Figures 6.3 & 6.4). For DF in in swing there are two graphs in each condition, in order to highlight differences among participants with and without participant 3 who was an outlier. From Figure 6.5 we can see that four out of seven pwMS had a higher direct orthotic effect of ankle angle at initial contact at the dual-task condition and only three out of seven after inducing fatigability. The individual graphs of the participants for the direct orthotic effect of walking speed showed that three out of seven pwMS had a higher effect during the dual-task condition and only for two out seven after completing the fatiguing task (Figure 6.6).

6.4.3 Dual-task cost of walking speed

Another aim of the present study was to investigate the DTC of walking speed (DTC_{WS}) without the use of FES (Task B) in pwMS and to compare it with an age-matched healthy control group. Both groups had a positive DTC_{WS}, with this positive percentage change indicating that there was a decrease in the walking speed from single to dual-task condition. However, it can be seen from the data in Table 6.5 that the DTC_{WS} in the MS was considerably higher compared to the healthy control group.

Table 6.5 Percentage DTC_{WS} of the MS and healthy control group (presented as means and SDs).

	Participant type	WS (m/s)	WS_{DT} (m/s)	DTC_{WS} (%)
	MS (n=7)	0.77 (0.4)	0.67 (0.4)	12.4 (11.5)
	HC (n=5)	1.26 (0.2)	1.20 (0.3)	4.6 (7.8)

6.4.4 Fatigability induced gait changes under dual-task conditions

Table 6.6 provides the medians (IQR) before and after completing the fatiguing task of the kinematic and spatiotemporal parameters of the two groups. It is apparent from this table that on average pwMS had slightly decreased peak DF in swing (1°) and ankle angle at initial contact (0.8°) after Task C compared to Task B. The healthy control group did not show any signs of fatigability since there was an increase on gait kinematics after exercise task. The healthy individuals had approximately ≈8° and ≈4° higher peak DF in swing and ankle angle at initial contact respectively compared to the MS group in both Task B and Task C. Further, the healthy control group walked faster, with increased cadence and step length in Task C and also compared to the MS group.

Table 6.6 Fatigability induced gait changes in kinematic and spatiotemporal parameters during dual tasking in pwMS (without FES) and healthy individuals (presented as medians and IQR).

	Task	MS (n=7)	HC (n=5)
DFMA (°)	$Unfatigued_{DT}$	1.3 (11.8)	9.0 (5.5)
	$Fatigued_{DT}$	0.3 (9.4)	9.2 (7.2)
ICMA (°)	$Unfatigued_{DT}$	1.1 (14.1)	3.7 (6.8)
	$Fatigued_{DT}$	0.3 (12.3)	3.9 (9.2)
WS (m/s)	$Unfatigued_{DT}$	0.76 (0.59)	1.10 (0.49)
	$Fatigued_{DT}$	0.57 (0.41)	1.24 (0.41)
Cadence (steps/min)	$Unfatigued_{DT}$	98 (67)	119 (25)
	$Fatigued_{DT}$	89 (48)	128 (22)
SLMA (m)	$Unfatigued_{DT}$	0.43 (0.09)	0.60 (0.15)
	$Fatigued_{DT}$	0.45 (0.17)	0.63 (0.12)

Abbreviations: DFMA: peak DF in swing of most affected leg; DT: dual-tasking; ICMA: ankle angle at initial contact of most affected leg; SLMA: step length of the most affected leg; WS: walking speed.

6.5 Discussion

6.5.1 Direct orthotic effect of FES under different conditions

The first aim of the present study was to investigate the direct orthotic effect of FES under different conditions, which are assumed to be more ecologically valid and would thus replicate the real life environment. Thus far, seven pwMS have been recruited in the study and have completed: walking only trials (Task A), walking trials while performing the Stroop test (Task B) and walking with the Stroop test after a task inducing fatigability (Task C).

Previous studies that have investigated the direct orthotic effect of FES on kinematic parameters, reported significant increases in peak DF in swing and at initial contact (Scott et al, 2013; van der Linden et al, 2014a; van der Linden et al, 2014b). Moreover, a recent meta-analysis reported a significant direct orthotic effect of FES on walking speed especially over short distance tests (Miller et al, 2017).

Even though there is a great amount of literature investigating the direct orthotic effect of FES either on kinematic or spatiotemporal parameters, the novel element of the present study is the exploration of the direct orthotic effect of FES in pwMS who are simultaneously performing a cognitive task. This is of clinical significance, as it has been reported by many pwMS that one of the benefits of FES is that less concentration is needed while walking and that it is easier to perform two tasks (e.g. walking and talking) at the same time.

Based on the qualitative studies by Bulley et al. (2015) and (Miller) Renfrew et al. (2018), we hypothesized that the direct orthotic effect of FES on ankle kinematics (peak DF in swing and ankle angle at IC) and spatiotemporal gait parameters would be higher in dual tasking conditions (Task B and Task C) compared to single walking

condition (Task A). For the peak DF in swing this was indeed the case for six out of seven people in the MS group, although for none of the participants was this increase higher than the $MDC_{95\%}$ (6.1°) derived from group B (EDSS range: 4-6) in Chapter 4 (reliability study). None of the participants exceeded the $MDC_{95\%}$ for initial contact (7.7°) reported in Chapter 4. For walking speed however, only three out of seven participants had a higher direct orthotic effect in both dual-task conditions (Task B and Task C) compared to the single task (Task A). A possible explanation for not finding a higher direct orthotic effect of FES in Task C compared to Task A in some of the participants might be that the action of FES was not sufficient to overcome the deficits due to local muscle fatigue.

These findings agree with only one other study that examined gait changes after a 6MWT and the effects of FES, which also reported that the action of FES alone was not sufficient to overcome the fatigability-induced gait changes after the fatiguing task (Barr et al, 2017). To the best of the author's knowledge, there are no published studies reporting on the efficacy of FES under dual-tasking conditions nor under dual-tasking combined with fatigability-induced conditions in pwMS or any other populations.

6.5.2 Dual-task cost of walking speed

Another aim of the present study was to investigate the DTC_{WS} in pwMS and a healthy control group. Our findings showed that the DTC_{WS} was higher in pwMS (12.4%) compared to the healthy individuals (4.6%), indicating that the decrement in walking speed from single to dual tasking was higher in pwMS compared to the healthy control group. Further, our hypothesis based on previous studies that the MS group would exhibit a DTC_{WS} of at least 10% was confirmed, since the DTC_{WS} in our MS group

was 12.4%. The percentage change in our MS group is in accordance with work of other studies in this area. One of the first studies exploring dual tasking in pwMS during "walking while talking", reported a DTC of walking speed of 10.7% for a titrated demand task and of 8.6% for a fixed demand task in a group of pwMS with mild to moderate disability level (Hamilton et al, 2009). In line with our findings, the majority of other studies exploring dual tasking in pwMS have reported a DTC of walking speed ranging from 11.8% to 15%. This is despite most of the studies employing a different cognitive task such as the modified word list generation and alternating letters of the alphabet (Sosnoff et al, 2011a; Learmonth et al, 2014; Motl et al, 2014; Coghe et al, 2018). Two studies also investigated the DTC of walking speed in pwMS and compared this with that in healthy individuals and reported higher DTC in pwMS (Hamilton et al, 2009; Coghe et al, 2018), similar to this study. Interestingly, the MS group in those two studies had mild disability level according to the EDSS, but they still reported statistically significant differences compared to healthy individuals. Despite the growing number of studies on DTC of gait in MS, there is no consensus on which cognitive task should be used to explore dual-task performance in pwMS. The selection of an appropriate cognitive task is important, since different cognitive tasks require different mental processes to execute them (Al-Yahya et al, 2011). A recent meta-analysis reported that the effect of cognitive-motor interference on motor performance was only significantly different between pwMS and healthy individuals during discrimination and decision-making tasks and not for example during mental tracking or verbal fluency tasks (Learmonth et al, 2017).

For that reason, we chose as a cognitive task the Stroop test, which is a discrimination and decision-making task and evaluates attention, working memory, processing speed

and cognitive flexibility (Stroop, 1935; Al Yahya et al, 2011). To the author's knowledge only one study investigated cognitive-motor interference (CMI) in spatiotemporal parameters with the Stroop test in pwMS with mild disability (EDSS mean = 2.1) and reported a DTC of walking speed of 11.7% which was significantly higher compared to a healthy control group (DTC = 4.8%) (Coghe et al, 2018). Interestingly, in their systematic review into CMI in pwMS, Learmonth et al. (2017) concluded that CMI does not differ in magnitude under single and dual task conditions between groups of pwMS and healthy individuals (ES =0.08). However, this absence of difference in CMI between pwMS and healthy individuals could be attributed to the inclusion of a variety of different cognitive tasks in that meta-analysis, since the individual studies described above did report statistically significant differences in those two groups (e.g. Hamilton et al, 2009; Wajda et al, 2013). As there is a growing interest for the investigation of CMI in pwMS, and the underlying mechanisms, future research should consider the use of DT paradigms for rehabilitation purposes that could potentially positively impact the quality of life of pwMS. Only one study by Peruzzi et al. (2016), reported improvement in walking speed and stride length in dual-task conditions after a virtual reality treadmill training program.

6.5.3 Fatigability induced gait changes under dual-task conditions

The third aim in the present study was to investigate gait characteristics after inducing fatigability between pwMS and healthy individuals while performing a cognitive task. Thus far, and to the best of our knowledge, no other studies have reported on gait changes after inducing fatigability and while simultaneously performing a cognitive task that potentially resembles conditions that pwMS experience in everyday life. A

study by Wolkorte et al. (2015) examined whether pwMS were more challenged by a dual-task compared to healthy individuals in fatiguing and less fatiguing conditions, but they were focused on maximal voluntary contractions of the index finger as a measure of fatigability and they did not examine gait changes.

The present study found a small average decrease in ankle peak DF in swing (-1.0º) and ankle angle at initial contact (0.8º) in pwMS in the trials after the fatigability-inducing task compared to those before. However, this decrease did not exceed the $MDC_{95\%}$ of group B (the most impaired group), presented in Chapter 4, in any of the participants. Similar to our findings, studies investigating fatigability-induced gait changes in pwMS with similar disability level to ours (EDSS range: 3-6), reported a small decrease in peak DF in swing and ankle angle at initial contact after a 6MWT (McLoughlin et al, 2016; van der Linden et al, 2018).

Interestingly, the healthy control group walked on average slightly faster (0.14 m/s), with associated increase in step length (0.03m) and cadence (9 steps/min) in the trials after the fatiguing task. This could possibly be explained by the post-activation potentiation effect (Sale, 2004). The healthy control group in Chapter 5 and the healthy control group in the study by McLoughlin et al. (2016) also presented the same phenomenon with increased walking speed, cadence and step length after an exercise task (i.e. 20 minute treadmill run and 6MWT respectively). However, the MS group did not present a similar increase in the spatiotemporal parameters, and it could possibly be explained by the disability of this group (EDSS 4-6).

6.5.4 Limitations

The present study has several limitations. Primarily, the study was under-powered. Further, our inclusion criteria were possibly not specific enough, for example we included a participant who had FES on both legs (Participant 6) and one who used FES only for running as she did not have foot drop during walking (Participant 5). This resulted in large between participant variability in the outcome measures, which reduced the effect size. Further, when performing the SRT to induce fatigability, participants could terminate the test at any point with all but one terminating the test approximately three to four minutes after the start. Even though most participants reported that they could not continue due to foot drop or lack of balance, the time performing the fatiguing task might not be long enough to observe greater differences between Task B and Task C (without FES). Finally, participants performed the Stroop test over a 7m distance on the gait laboratory, which could have resulted in a learning effect of the cognitive task while performing 24 walking trials in total. We have tried to minimize this possible effect by counterbalancing the walking trials with and without the Stroop test before the fatiguing task, but in all trials after the fatiguing task the Stroop test was performed.

6.6 Conclusion

The present study set out to investigate the direct orthotic effect of FES on gait characteristics under a variety of walking conditions, such as walking while performing a cognitive task and walking after inducing fatigability with a simultaneous cognitive task. To the author's knowledge, this is the first study to attempt to compare the direct orthotic effect of FES in different walking conditions in pwMS or any other neurological condition.

Secondary objectives were to examine the difference in DTC_{WS} and fatigability measured as the deterioration in gait characteristics between pwMS and a healthy control group. Individual results indicated that for most participants the direct orthotic effect of FES was higher under dual-task conditions. However, the effect sizes for peak DF in swing, ankle angle at IC and walking speed between Task B and Task A were small (d=0.12; d=0.04; d=0.26 respectively). Similarly, small but positive effect size (d=0.1) was found for the comparison of the direct orthotic effect on peak DF in swing between Task C and Task A.

Our hypothesis that the DTC_{WS} for pwMS would be at least 10% was supported by our findings (DTC_{WS} = 12.4%) and this was clearly higher than that found for the healthy control group (4.6%). After inducing fatigability there was a small decrease in the kinematic parameters and spatiotemporal parameters for the pwMS. Notwithstanding the relatively limited sample, this study offers valuable insights into the benefits of FES that pwMS report. In particular, that FES reduces the mental effort of walking. Further studies need to be carried out to investigate the benefits of FES under dual-task conditions that will include a more homogeneous sample of pwMS and possibly under more demanding dual-task conditions. Further, more research is needed to examine dual-task performance in pwMS with the use of a standardized cognitive task in order to compare findings amongst studies.

Chapter 7. Final discussion

7.1 Purpose of the chapter

The purpose of this final chapter is to synthesize the findings from the studies included in this thesis in an integrated discussion of the results and limitations. In the previous chapters, each study was discussed independently, but this final discussion will attempt to synthesize the findings in a broader context of existing literature.

Firstly, a summary of all the key findings of each of the studies will be presented (section 7.2), followed by a synthesis of the findings in relation to the main focus of the present PhD that is investigation of walking impairments and especially foot drop in pwMS. Finally, methodological considerations, recommendations for future work and clinical implications will be presented throughout the chapter.

7.2 Summary of the findings and linkage between the chapters

The key findings from the chapters in this thesis are summarised in Table 7.1. The purpose of the systematic review (Chapter 3) was to evaluate the psychometric properties of outcomes measures used to examine the effect of assistive technology to treat foot drop in pwMS. One of the findings was the lack of psychometric evidence of 3D gait kinematics in the MS population, even though it was one of the most used outcome measures to evaluate walking performance with and without assistive technology. Thus in the next chapter (Chapter 4), the test-retest reliability of ankle kinematics and spatiotemporal parameters in two groups of pwMS with different level of walking impairment was examined. The SEM and $MDC_{95\%}$ indices for ankle

kinematics derived from this study were then used to interpret the results of the study included in Chapter 5. This chapter focused on the objective measurement of fatigable foot drop (i.e. foot drop that recovers with rest) that is subjectively reported by highly physically active pwMS whose daily walking ability is not affected. The indices derived in Chapter 4 were also applied to interpret the results of the last study (Chapter 6) that focused on a group of pwMS that had established (i.e. 'fixed') foot drop and were treated for this with FES. The main focus of this last study was to investigate the direct orthotic effect of FES under a variety of walking conditions, such as dual-tasking and exercise-induced fatigability, that were regarded more ecologically valid than short single task walking tests commonly used to assess the effect of FES.

Table 7.1 Key findings of the four studies presented in this thesis.

Chapter	Key findings
Chapter 3: Walking measures to evaluate assistive technology for foot drop in multiple sclerosis: A systematic review of psychometric properties.	• Most frequently used outcomes measures to evaluate walking performance were walking speed (mostly recorded over 10m distance), 3D gait kinematics and the Physiological Cost Index. • No psychometric studies for 3D gait kinematics & PCI in MS population. • Moderate to strong evidence for the psychometric properties for MSWS-12, T25FW, 6MWT & 10mWT.
Chapter 4: Test-retest reliability and minimal detectable change of ankle kinematics and spatiotemporal parameters in pwMS with low (< 3.5) and moderate to high EDSS (4-6).	• Reliability was excellent (ICC values > 0.75) for ankle kinematics and spatiotemporal parameters. • GPS presented fair to moderate ICC values in both groups. • Smaller $MDC_{95\%}$ values in the low EDSS group compared to the higher EDSS group.

Chapter 5: Deterioration of gait characteristics in minimally impaired pwMS after an exercise task.	• Significant decrease in peak DF in swing for pwMS but no change in spatiotemporal parameters after a 20min run on a treadmill at a self-selected speed compared to the healthy control group. • Six out of the 15 pwMS had a decrease in peak DF in swing exceeding the $MDC_{95\%}$ value, whilst no one on the healthy control group showed such a decrease. • Exercise-induced decrease in DF suggests a sign of onset of MS related fatigability.
Chapter 6: The direct orthotic effect of functional electrical stimulation on gait kinematics and walking speed in people with	• Walking with the assistance of FES resulted in an improvement in all the gait parameters and especially for peak DF in swing. • There was a small trend towards a higher direct orthotic effect for peak DF in swing during dual-task conditions (B and C) compared to single task (A) (ES Task A vs B = 0.12, ES Task A vs C = 0.10).

MS under dual-tasking and fatiguing walking conditions simulating daily life.	• The DTC_{WS} was higher in the MS (12.4%) compared to the healthy group (4.6%). • There was a decrease in peak DF in swing and AAIC in pwMS after inducing fatigability without the use of FES.

7.3 Psychometric properties and outcome measures of walking performance

7.3.1 Reporting standards on psychometric properties studies

There are many guidelines to promote a robust study design and high quality reporting standards. For example, for randomised trials, the Consolidated Standards of Reporting Trials (CONSORT) aims to standardise the reporting of findings and to facilitate a transparent and complete reporting of the study methodology of randomised controlled trials (Schulz et al, 2010). The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) provides a guide to develop protocols for systematic review and meta-analyses (Moher et al, 2010). In Chapter 3, the systematic review was based on the COSMIN guidelines, which aim to provide an overview of the quality of an outcome measure (i.e. its psychometric properties) and support evidence-based recommendations for the selection of the most appropriate outcome measure for a specific purpose (Mokkink et al, 2010c). The COSMIN checklist was originally developed to rate the methodological quality of patient-reported outcome measures. However, the COSMIN developers have stated that systematic reviews on performance based measures can still be conducted based on the COSMIN guidelines, but some of the items on the checklist would need to be adapted to fit the reporting of the methodology (Mokkink et al, 2012). The COSMIN guidelines were updated on 2018 after we published our systematic review. One of the major changes that are included in the new guidelines is that studies rated as 'poor' (i.e. in the new guidelines are termed as 'inadequate') are no longer omitted from the rating of evidence. Previously, the 'poor/inadequate' studies were not taken into account in the evaluation of the overall quality of an outcome measure because the results of these studies

could be subject to bias. However, in the new guidelines these studies are included with the reasoning that all studies can be included but the quality of the evidence can be downgraded in the risk of bias assessment (Prinsen et al, 2018). Another major change is that items with regard to formulating hypotheses *a priori* for 'hypothesis testing' (i.e. construct validity) and responsiveness are removed, since it was concluded that the results from studies not reporting this information can still report appropriate correlations and the results are not necessarily biased. Finally, the item of adequate sample size has been removed from the boxes evaluating each measurement property to a later stage of the review process, because it is believed by the authors of the guidelines that even small but high quality studies can provide appropriate evidence of the measurement property evaluated. It is acknowledged that these changes of the guidelines could have impacted on our already published findings and conclusions. For example, a few studies in our systematic review were rated as 'poor/inadequate' due to either small sample size or because *a priori* hypotheses were not formulated. However, the overall level of evidence for each outcome measure would not be greatly affected, since the results of each study were also evaluated separately by the criteria for good measurement properties and were combined with the methodological quality to obtain an overall level of evidence for each outcome measure.

The COSMIN taxonomy has been criticised for not including a domain for reliability of change score (i.e. 'longitudinal reliability') and for not supporting the use of MDC as an index for the domain of reliability of change scores (Polit, 2016). However, the developers of COSMIN strongly support the use of MDC only as an index of measurement error, since it is calculated by the SEM (i.e. a parameter of measurement error) and this is

rarely based on an empirical assessment of the reliability of the change score (Mokkink et al, 2016). The COSMIN group continues to update the COSMIN tools, with the revised guidelines including a ten step process of performing a systematic review based on standards (i.e. referring to the design and statistical methods for evaluating the quality of studies on measurement properties) and criteria (i.e. what constitutes good measurement properties) (Prinsen et al, 2018).

7.3.2 Evidence for psychometric properties of measures of walking performance in pwMS

Walking ability is affected in many pwMS and there is a wide range of outcome measures evaluating walking in pwMS. In Chapter 3, 20 outcome measures were identified assessing walking performance, lower limb function and effort of walking in studies evaluating the effect of assistive technology in pwMS. However, only ten of these outcome measures were evaluated for their psychometric properties. This is of concern as it is essential for both clinicians and researchers to choose the most appropriate outcome in order to evaluate the effectiveness of an intervention such as FES to treat foot drop, which is the focus of this thesis. The selection of such an outcome should be based on its psychometric properties. By critically evaluating the evidence of studies on measurement properties of these walking performance outcomes using the COSMIN guidelines, we found strong levels of evidence for the internal consistency and test-retest reliability of the MSWS-12 and the construct validity of the T25FW. Moderate levels of evidence were found for test-retest reliability and responsiveness of both the T25FW and 6MWT. Interestingly, the 10mWT was the most frequent outcome measure used but only a

moderate level of evidence was found for responsiveness. However, this finding does not necessarily indicate that the 10mWT should not be used in clinical and research settings since it could have resulted from 'poor' methodological quality studies. However, this highlights the need for better design and planning of psychometric studies. Thus far, only two systematic reviews have reported on the outcome measures for the evaluation of walking performance in pwMS. Kieseier and Pozzilli (2012) reported on the psychometric properties of walking performance measures, but they did not assess the methodological quality of the included studies. The outcome measures were categorised as those assessing walking speed (e.g. T25FW, 10mWT, 30mWT), those evaluating walking distance (e.g. 2MWT and 6MWT) and self-reported measures such as the MSWS-12. It was recommended that the T25FW was the most well characterised objective measure to be used in clinical settings, whilst longer timed or distance measures should be used to assess fatigability, distance limitations and functional capacity (Kieseier & Pozzilli, 2012). Bethoux and Bennett (2011) also recommended the T25FW along with the MSWS-12 to be suitable for routine evaluation of walking performance in the MS population due to its psychometric properties. However, the methodological quality of the studies included in these reviews was not taken into account in these recommendations. It is of importance to conduct high quality studies evaluating the psychometric properties of these outcomes. In particular, there is a need for studies providing indices like the minimal detectable change and minimal clinically important difference, both of which provide essential information to clinicians and researchers when interpreting score changes after an intervention.

Even though in Chapter 3 it was highlighted that no studies had evaluated the psychometric properties of 3DGA in the MS population, a topical review emphasised its

importance for quantifying gait abnormalities, which could potentially allow for a better understanding of the underlying mechanisms of walking impairments that measures like spatiotemporal parameters cannot detect (Cofré Lizama et al, 2016). In order to address some of the gaps in the evidence for the measurement properties of 3DGA identified in Chapter 3, a test-retest reliability (i.e. absolute and relative reliability) study was performed in this thesis (Chapter 4). The study was focused on test-retest reliability of spatiotemporal parameters and especially ankle kinematics of pwMS, since foot drop is quantified by the decrease in dorsiflexion/increase in plantarflexion. Further, this study provided indices like SEM and $MDC_{95\%}$ which provide essential information to clinicians and researchers. The test-retest reliability study included two distinct groups of pwMS (Group A: EDSS 0-3.5; Group B: EDSS 4-6). Participants in Group A were both minimally impaired by MS and highly physically active while participants in Group B used either an AFO or FES to treat their foot drop. Good to excellent ICC values for all kinematic and spatiotemporal parameters were found in both groups, with the SEM and $MDC_{95\%}$ being lower for Group A compared to Group B. A higher variability in walking performance in pwMS with a higher level of walking impairment has been observed in other studies. For example, it has been shown that pwMS with EDSS > 4.5 and using assistive devices have great variability in spatiotemporal parameters than those who do not (Socie et al, 2013; Socie et al, 2014).

Three-dimensional gait analysis systems utilise multiple cameras to track three-dimensional trajectories of markers placed on anatomical landmarks (e.g. Vicon, Optotrak, etc.). These systems are considered the ‘gold’ standard for recording gait kinematics and 3DGA is often utilised to establish the construct validity of other measures

such as the Microsoft Kinect system that has emerged as a tool for movement analysis in clinical practice (Asaeda et al, 2018). However, Chapter 3 showed that EDSS is often used as the gold standard (or comparator) when assessing the criterion (or construct validity) of outcome measures used to assess walking performance. The EDSS is a scale that evaluates general disability and walking performance is taken into account only at EDSS of 4.0 and higher, which is important to bear in mind when the population of interest includes people with EDSS scores less than 4.0.

Three-dimensional gait analysis outcomes and especially indices, which quantify the 'overall' deficit in gait kinematics, such as the Gait Profile Score may be more appropriate comparators for clinical outcomes of walking performance. Although the reliability of the gait kinematic parameters has been assessed in other populations (e.g. Redekop et al, 2008; Devetak et al, 2016), the study reported in Chapter 4 is the first to provide evidence of the reliability of gait kinematics in pwMS. Further work is needed to establish the psychometric properties of gait kinematics in pwMS other than reliability, such as the minimal clinically important difference (MCID). Knowledge of the MCID would assist the interpretation of studies exploring the long-term change in gait kinematics or those investigating the long-term effect of FES to treat foot drop.

7.4 Gait deterioration and foot drop

7.4.1 Foot drop in minimally impaired pwMS

Consultant neurologists at the Anne Rowling Clinic in Edinburgh observed that it is not uncommon for highly physically active pwMS with EDSS < 2.5 to report that they

experience symptoms such as foot drop during relatively high intensity and/or long duration exercise bouts. These symptoms may affect walking ability during and directly after exercise but will disappear after a relatively short period of rest. Following these reports and reports of 'fatigable foot drop' in people who use FES (Bulley et al, 2015), we conducted an audit at the Anne Rowling Clinic which showed that 70% pwMS (33 out of 47 respondents) reported experiencing transient foot drop that recovered following rest. Moreover, it was indicated from this audit that the most commonly reported activities that seem to induce this phenomenon were walking and running, while the average time until the initiation was 20 min (range: 2-60 min) and average distance covered until initiation was 1.6 miles (range: 0.5-4.5 miles). Thus, the aim of Chapter 5 was to attempt to objectively quantify this exercise-induced foot drop in highly physically active and minimally impaired (EDSS < 2.5) pwMS. The findings of the audit allowed us to hypothesize that for most participants a 20 min treadmill run at a self-selected speed would result into exercise-induced foot drop. Indeed, in the study described in Chapter 5 it was found that there was a statistically significant decrease in peak DF in swing in pwMS compared to healthy individuals after a 20-minute treadmill run, with six out of the 15 pwMS showing a decrease in peak DF in swing exceeding the $MDC_{95\%}$ (as derived in Chapter 4). However, it was observed that some of the participants who reported experiencing foot drop did not show any significant changes in kinematic parameters. A possible explanation for this might be that the test protocol used in this study was not demanding enough (i.e. too short or low intensity exercise task) for this highly physically active group of pwMS.

To our knowledge, the study described in Chapter 5 is the first study that investigated gait kinematics and spatiotemporal parameters after an exercise task that attempted to induce fatigability in highly physically active pwMS who did not experience any habitual gait dysfunction. Only a few studies have investigated changes in gait kinematics after inducing fatigability (i.e. after a 6MWT or walking on a treadmill until exhaustion), but the MS population in those studies had higher EDSS (EDSS > 3.5) (Sehle et al, 2011; McLoughlin et al, 2016).

Another interesting finding of the study in Chapter 5 was that the pre-exercise gait characteristics in the highly physically active participants with MS did not differ from the age-matched healthy control group, contrary to findings in previous studies with pwMS with low EDSS. It is possible that the fact that the participants were all highly physically active could accounts for this difference in findings. Reporting the physical activity status of participants in studies investigating gait characteristics in pwMS with minimal or no walking impairment in daily life, would enhance the interpretation of the findings. Future studies should investigate transient foot drop in pwMS with EDSS scores < 3.5 but who are not highly physically active and thus may be a more representative sample. Further, this supports the notion that a more demanding task, such as the protocol used in this study, is needed in order to highlight changes in minimally impaired pwMS.

In Chapter 5, fatigability was quantified as exercise-induced foot drop. A recent systematic review summarised existing protocols and outcome measures used to detect fatigability in the MS population (Severijns et al, 2017). It was found that most protocols used maximal single-joint isometric contractions, whilst walking protocols included the 6MWT, walking on a treadmill until exhaustion and the T500mW (Severijns et al, 2017).

The majority of the studies, which used walking protocols to induce fatigability, quantified fatigability as the decrease in walking speed over time. However, a decrease in walking speed can also be the result of non-MS specific factors such as pacing and aerobic capacity (Dalgas et al, 2014). This study together with studies by McLoughlin et al. (2016) and Sehle et al. (2014) instead focused on gait kinematics where are more likely to reflect underlying MS specific neurological deficits. However, there is no standardised protocol in order to compare findings among studies, since each protocol is measuring different aspects of fatigability and future work should focus on investigating the psychometric properties of outcome measures used to assess fatigability in pwMS.

Fatigability has also been investigated in different populations and similarly there are no standard protocols to assess fatigability in these populations either. For example, fatigability has been documented as changes in kinematic and spatiotemporal parameters after nine consecutive trials (over 10m) in stroke survivors (Boudarham et al, 2013), in myasthenia gravis as the distance covered in six 60s intervals of the 6MWT (Jordan et al, 2017) and in older women as a decrease of maximum velocity of consecutive repetitions of sit-to-stand transfers (Lindemann et al, 2016).

A potential factor that could have influenced the findings with regard to the presence of exercise-induced foot drop is the level and number of spinal lesions and the atrophy of the spinal cord in our MS group. A recent study by Sechi et al. (2019) reported that for pwMS with unilateral motor progression, the motor deficit may be attributable to a single critical corticospinal tract lesion. To the author's knowledge, the correlation of spinal lesions with functional and walking performance outcome measures has not been widely examined. Only one study, by Cohen et al. (2012), reported that the upper cervical spinal cord volume

significantly correlated with EDSS, while there was lack of correlation between all brain and spinal cord lesions measures and the T25FW in a group of mildly disabled MS group with RRMS. In this case this might be explained by the fact that this group has no walking impairments. Another study also reported that the atrophy of the spinal cord is related to clinical disability as assessed with the EDSS (Lukas et al, 2013). However, the location and the number of lesions in the participants in the study were not investigated in Chapter 5 as it was out with the scope of the overall aim of this thesis. However, the association between spinal lesions and walking related symptoms such as transient foot drop requires further investigation in order to gain an improved understanding of the underlying mechanisms of this exercise-induced phenomenon. Further work that could identify potential mechanisms of fatigability induced gait changes is through electromyography (EMG) for example of the tibialis anterior and the gastrocnemius muscles.

7.4.2 The effect of FES to treat foot drop and dual-tasking in people with MS whose walking ability is impaired by fixed foot drop

Even though there is a great amount of literature investigating the direct orthotic effect of FES on kinematic and/or spatiotemporal parameters, and usually over short distance/duration, the novel element of the study described in Chapter 6 is the exploration of the direct orthotic effect of FES on walking characteristics in pwMS who are simultaneously performing a cognitive task.

Chapter 6 aimed to investigate the direct orthotic effect of FES under a variety of walking conditions that would be more ecologically valid and thus replicate activities of daily life in comparison to standard single task walking tests commonly used in research and

clinical practice. Participants were asked to perform walking trials under three conditions: 1) single task walking, 2) walking with the addition of a cognitive task and 3) walking after inducing fatigability while simultaneously performing a cognitive task. The idea of this study was based on recent evidence of the experiences of the use of FES which suggests that one of the key perspectives that pwMS express is that they 'do not have to concentrate as hard or think about every step they were taking' when using FES (Bulley et al, 2015; (Miller) Renfrew et al, 2018). Even though there is a growing body of evidence on the effect of dual-tasking on gait parameters in pwMS, to our knowledge no other study has objectively examined and reported on the benefits of FES use under dual-tasking conditions. The results indicated that there was a small trend for higher direct orthotic effect of FES in peak DF in swing during the dual-task conditions (i.e. dual-task and dual-task after induced fatigability) compared to single task condition (i.e. only walking) with a small but positive effect size (≈0.12). The DTC of walking speed in the condition without induced fatigue was 12.4%, while for the healthy control group was 4.6% which is similar to that found in previous studies investigating the DTC of walking speed compared to healthy individuals (e.g. Sosnoff et al, 2011a; Learmonth et al, 2014; Coghe et al, 2018). The sample size of the present study was less than required for an appropriately powered trial and thus we were not able to perform any inferential statistics. Hence, further appropriately powered trials are needed to confirm these findings. However, this study showed that the protocol was feasible in this population of pwMS and the results provide an initial insight into the potential benefits of FES which have not been explored before.

There is evidence for the interaction of cognitive and motor functions in MS, with a growing interest in CMI that is the decline in the performance on the motor task while

performing a cognitive task simultaneously (Benedict et al, 2011; Motl et al, 2016). Nonetheless, a recent systematic review reported equivocal findings of studies investigating the DTC of walking speed and concluded that there is a minimal difference in CMI between pwMS and healthy individuals (Learmonth et al, 2017). It should be noted though, that the majority of the included studies had small sample sizes and more importantly used different cognitive tasks which makes it difficult to draw any conclusions since different test protocols examine different cognitive domains. The meta-analysis by Learmonth et al. (2017) suggested that discrimination and decision-making tasks to challenge motor control are most likely to highlight potential differences in CMI in pwMS compared to healthy individuals. For that reason, we decided to use the Stroop test as a cognitive task which is a discrimination and decision making task and is associated with the measurement of attention and response inhibition (MacLeod, 1991). Thus far, only one study investigated CMI on spatiotemporal parameters in pwMS with the use of the Stroop test and reported a DTC of walking speed of 11.7% for pwMS and 4.8% for the healthy control group (Coghe et al, 2018). Although similar to our findings, it should be noted that the MS group in that study consisted of mildly disabled pwMS (EDSS mean: 2.1) whereas our group were moderately disabled (EDSS range: 4.0-6.0), suggesting that CMI is evident in mild MS compared to healthy individuals. Future studies should explore whether there is a correlation between CMI and disability status.

The importance of dual-tasking in pwMS was also highlighted in a review by Motl et al. (2016) which suggested dual-task exercises as a promising intervention for improving walking and cognitive functions in pwMS. Although there is a growing body of literature highlighting the importance of exercise and cognitive interventions for rehabilitation

purposes in pwMS, the authors of this review concluded that there are no studies examining the efficacy of combining these two types of interventions on walking performance in pwMS, for example through dual-task exercises (Motl et al, 2016). Hence, the individual and combined effects of exercise and cognitive rehabilitation and their interaction on walking and cognitive performance need to be explored in future studies.

7.5 Limitations and future work

The present thesis has several limitations that should be considered in future studies. Firstly, the systematic review in Chapter 3 that reported on the evidence of the psychometric properties of self-reported and objective outcome measures of walking performance was based on the COSMIN guidelines that were developed for HR-PROMs and not for the type of performance-based measures included in the review in Chapter 3. However, the theoretical framework underpinning the development of COSMIN was applicable to both HR-PROMs and performance based outcome measures and the updated guidelines support the use of COSMIN on other measures apart from HR-PROMs with adapted methodology.

Chapter 4 examined the test-retest reliability of ankle kinematics and spatiotemporal parameters in two groups of pwMS with different level of walking impairments. The focus of the study was on the outcomes that can evaluate and quantify foot drop and that was the reason we examined test-retest reliability of the ankle kinematics. Nonetheless, and since they can provide important information of compensation strategies for foot drop in this population, future research should be undertaken to examine the reliability of the

kinematic parameters in other joints of the lower limbs (i.e. knee, hip and pelvis). Moreover, responsiveness of gait kinematics should also be examined since it is essential information for evaluating meaningful changes both in routine clinical practice and research interventions. Further, the sample size in both groups of pwMS would be considered small according to the COSMIN guidelines. However, in the updated guidelines the developers of COSMIN have removed the standard of 'adequate' sample size, since it was decided that small high-quality studies can still provide sufficient information on a measurement property and it is taken into account in a later phase when conclusions on the overall level of evidence of an outcome measure are drawn.

The main weakness of the study in Chapter 5 was the exercise task that consisted of a 20 minutes self-selected speed run on a treadmill. This exercise task was chosen based on the finding in our foot drop audit (see Appendix 1) that the average time until the onset of transient foot drop was 20 minutes. However, the chosen exercise task in this study might not have been demanding enough for some of the participants, since we were not able to capture exercise-induced foot drop for some participants who reported experiencing this phenomenon. Another limitation is that the use of a treadmill limits the ecological validity and the clinical implementability of that specific exercise task. To develop a more comprehensive insight into fatigability in mildly disabled pwMS, additional studies will be needed. Those studies should consider the use of standardised clinical tests based on each participants' functional ability such as the shuttle walk test, which is incremental and progressive and stresses the individuals to their maximal performance (Singh et al, 1992).

The main limitation in Chapter 6 was the small number of participants in both MS and healthy control groups and thus the study was not powered to detect any statistically

significant differences between the different conditions. The aim is to continue with recruitment until we reach the sample size derived from the a priori power calculation that was performed for this study. Further, because of the relatively broad inclusion criteria, a large variability among the participants was observed. For example, the population in the study consisted of a person using FES in both limbs and a person using FES only for running. Stricter inclusion/exclusion criteria should be applied in future studies regarding the use of FES. Further work is required to gain a better understanding of CMI in pwMS and the perceived benefits of FES and this can be achieved by standardised use of cognitive tasks evaluating the same cognitive processes (i.e. discrimination and decision making tasks) so that the findings of studies can be comparable.

Chapter 8. Conclusion

Walking impairments is one of the most debilitating symptoms in MS and especially foot drop although interestingly, its prevalence has not been reported in the literature. The overall aim of the present PhD thesis was to explore foot drop, its presence in pwMS with different disability levels and the psychometric properties of outcomes used to evaluate walking impairments.

The systematic review summarised all self-reported and objective measures of walking performance used in studies that evaluated the effects of assistive technology in pwMS and evaluated the level of evidence for their psychometric properties. The findings of this review can guide clinicians and researchers to choose the most appropriate outcome measure for their needs based on the level of evidence of each outcome. This review also highlighted that there was a gap in the literature, with regard to the psychometric properties of 3D gait kinematics in pwMS. This finding resulted in our next study examining the test-retest reliability of ankle kinematics and spatiotemporal parameters and in two groups of pwMS with different levels of walking impairments. The findings of this reliability study helped to interpret the results in Chapter 5, which objectively documented gait deterioration evidenced as foot drop induced by an exercise task in highly active pwMS that could potentially be a sign of onset of fatigability. It was also used in Chapter 6 after inducing fatigability in pwMS that were FES users. Future studies are needed to explore the exercise-induced gait deterioration over time using more standardised exercise tasks so that participants are performing the same test each time without the issues of self-selecting speed and training effects. Further exploration is also

needed to unveil the potential value of gait analysis as a sensitive tool for disease progression. Gait analysis can record minimal changes and longitudinal follow-up of pwMS can reveal potential deterioration of mobility in early stages of the disease process, however its responsiveness has not been explored for pwMS. Finally, the last study (Chapter 6) was the first to investigate the benefits of FES under dual-tasking and fatiguing conditions in pwMS. Although the study was based on a small sample of participants, the findings suggest that there was a positive trend for the direct orthotic effect of FES under dual-tasking conditions compared to single-task conditions. Further investigations with appropriately powered study designs are required to determine the effectiveness of FES and also the CMI in this population with standardised cognitive tests that assess the same cognitive processes.

In summary, the studies presented in this PhD have attempted to answer questions related to the level of psychometric evidence of walking performance measures in the MS population, as well as investigating foot drop in pwMS with different disability status. There are still many unanswered questions that the present thesis could not address, such as the underlying mechanisms of fatigability induced foot drop or the mechanisms of CMI in pwMS, but the project hopefully has shed some light on some of the issues and may ultimately positively impact the lives of pwMS.

References

(2005). The Goldman Consensus statement on depression in multiple sclerosis. *Multiple Sclerosis*, 11(3): 328-337.

Al-Yahya, E., Dawes, H., Smith, L., Dennis, A., Howells, K., Cockburn, J., (2011). Cognitive motor interference while walking: a systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 35(3): 715-728.

Albrecht, H., Wötzel, C., Erasmus, L.P., Kleinpeter M., König, N., Pöllmann, W., (2001). Day-to-day variability of maximum walking distance in MS patients can mislead to relevant changes in the Expanded Disability Status Scale (EDSS): average walking speed is a more constant parameter. *Multiple Sclerosis*, 7(2): 105-109.

Allali, G., Laidet, M., Armand, S., Elsworth-Edelsten, C., Assal, F., Lalive, P.H., (2016). Stride time variability as a marker for higher level of gait control in multiple sclerosis: its association with fear of falling. *Journal of Neural Transmission (Vienna)*, 123(6): 595-599.

Amatya, B., Young, J., Khan, F., (2018). Non-pharmacological interventions for chronic pain in multiple sclerosis. *Cochrane Database of Systematic Reviews*, 12: Cd012622.

Ammann-Reiffer, C., Bastiaenen, C.H., de Bie, R.A., van Hedel, H.J., (2014). Measurement properties of gait-related outcomes in youth with neuromuscular diagnoses: a systematic review. *Physical Therapy*, 94(8): 1067-1082.

Andreasen, A. K., Jakobsen, J., Petersen, T., Andersen, H., (2009). Fatigued patients with multiple sclerosis have impaired central muscle activation. *Multiple Sclerosis*, 15(7): 818-827.

Andreopoulou, G., Mercer, T.H., van der Linden, M.L., (2018). Walking measures to evaluate assistive technology for foot drop in multiple sclerosis: A systematic review of psychometric properties. *Gait & Posture*, 61: 55-66.

Armutlu, K., Korkmaz, N.C., Keser, I., Sumbuloglu, V., Akbiyik, D.I., Guney, Z., Karabudak, R., (2007). The validity and reliability of the Fatigue Severity Scale in Turkish multiple sclerosis patients. *International Journal of Rehabilitation Research*, 30(1): 81-85.

Arnett, P. A., Higginson, C.I., Voss, W.D., Bender, W.I., Wurst, J.M., Tippin, J.M., (1999). Depression in multiple sclerosis: relationship to working memory capacity. *Neuropsychology*, 13(4): 546-556.

Asaeda, M., Kuwahara, W., Fujita, N., Yamasaki, T., Adachi, N., (2018). Validity of motion analysis using the Kinect system to evaluate single leg stance in patients with hip disorders. *Gait & Posture*, 62: 458-462.

Baert, I., Freeman, J., Smedal, T., Dalgas, U., Romberg, A., Kalron, A., et al. (2014). Responsiveness and clinically meaningful improvement, according to disability level, of five walking measures after rehabilitation in multiple sclerosis: a European multicenter study. *Neurorehabilitation Neural Repair***,** 28(7): 621-631.

Baker, R., McGinley, J.L., Schwartz, M.H., Beynon, S., Rozumalski, A., Graham, H.K., Tirosh, O., (2009). The gait profile score and movement analysis profile. *Gait & Posture*, 30(3): 265-269.

Bakshi, R. (2003). Fatigue associated with multiple sclerosis: diagnosis, impact and management. *Multiple Sclerosis*, 9(3): 219-227.

Balantrapu, S., Sosnoff, J.J., Pula, J.H., Sandroff, B.M., Motl, R.W., (2014). Leg spasticity and ambulation in multiple sclerosis. *Multiple Sclerosis International*, 2014: 649390. doi: 10.1155/2014/649390.

Barr, C. J., Patritti, B.L., Bowes, R., Crotty, M., McLoughlin, J.V., (2017). Orthotic and therapeutic effect of functional electrical stimulation on fatigue induced gait patterns in people with multiple sclerosis. *Disability and Rehabilitation: Assistive Technology*, 12(6): 560-572.

Barrett, C. and P. Taylor, (2010). The effects of the odstock drop foot stimulator on perceived quality of life for people with stroke and multiple sclerosis. *Neuromodulation*, 13(1): 58-64.

Barrett, C. L., Mann, G.E., Taylor, P.N., Strike, P., (2009). A randomized trial to investigate the effects of functional electrical stimulation and therapeutic exercise on walking performance for people with multiple sclerosis. *Multiple Sclerosis*, 15(4): 493-504.

Bates, A. V., McGregor, A.H., Alexander, C.M., (2016). Reliability and minimal detectable change of gait kinematics in people who are hypermobile. *Gait & Posture,* 44: 37-42.

Beal, C. C., Stuifbergen, A.K., Brown, A., (2007). Depression in multiple sclerosis: a longitudinal analysis. *Archives of Psychiatric Nursing,* 21(4): 181-191.

Beckerman, H., Roebroeck, M.E., Lankhorst, G.J., Becher, J.G., Bezemer, P.D., Verbeek, A.L., (2001). Smallest real difference, a link between reproducibility and responsiveness. *Quality of Life Research,* 10(7): 571-578.

Benedetti, M. G., Piperno, R., Simoncini, L., Bonato, P., Tonini, A., Giannini, S., (1999). Gait abnormalities in minimally impaired multiple sclerosis patients. *Multiple Sclerosis,* 5(5): 363-368.

Benedict, R. H., Cookfair, D., Gavett, R., Gunther, M., Munschauer, F., Garg, N., Weinstock-Guttman, B., (2006). Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *Journal of the International Neuropsychological Society,* 12(4): 549-558.

Benedict, R. H., Munschauer, F., Linn, R., Miller, C., Murphy, E., Foley, F., Jacobs, L., (2003). Screening for multiple sclerosis cognitive impairment using a self-administered 15-item questionnaire. *Multiple Sclerosis,* 9(1): 95-101.

Benedict, R. H. and Zivadinov, R., (2011). Risk factors for and management of cognitive dysfunction in multiple sclerosis. *Nature Reviews Neurology,* 7(6): 332-342.

Bethoux, F. and Bennett, S., (2011). Evaluating walking in patients with multiple sclerosis: which assessment tools are useful in clinical practice? *International Journal of MS Care,* 13(1): 4-14.

Betts, C. D., Jones, S.J., Fowler, C.G., Fowler, C.J., (1994). Erectile dysfunction in multiple sclerosis. Associated neurological and neurophysiological deficits, and treatment of the condition. *Brain,* 117 (Pt 6): 1303-1310.

Bland, J. M. and Altman, D. G., (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research,* 8(2): 135-160.

Borg, G., (1982). Ratings of perceived exertion and heart rates during short-term cycle exercise and their use in a new cycling strength test. *International Journal of Sports Medicine,* 3(3): 153-158.

Boudarham, J., Roche, N., Pradon, D., Bonnyaud, C., Bensmail, D., Zory, R., (2013). Variations in kinematics during clinical gait analysis in stroke patients. *PLoS One,* 8(6): e66421. doi: 10.1371/journal.pone.0066421.

Browne, P., Chandraratna, D., Angood, C., Tremlett, H., Baker, C., Taylor, B.V., Thompson, A.J., (2014). Atlas of Multiple Sclerosis 2013: A growing global problem with widespread inequity. *Neurology,* 83(11): 1022-1024.

Brownlee, W. J., Altmann, D.R., Alves da Mota, P., Swanton, J.K., Miszkiel, K.A.,Wheeler-Kingshott, C.G., et al. (2017). Association of asymptomatic spinal cord lesions and atrophy with disability 5 years after a clinically isolated syndrome. *Multiple Sclerosis,* 23(5): 665-674.

Bulley, C., Mercer, T.H., Hooper, J.E., Cowan, P., Scott, S., van der Linden, M.L., (2015). Experiences of functional electrical stimulation (FES) and ankle foot orthoses (AFOs) for foot-drop in people with multiple sclerosis. *Disability Rehabilitation: Assistive Technology,* 10(6): 458-467.

Burridge, J., Taylor, P.N., Hagan, S., Swain, I., (1997). Experience of clinical use of the Odstock dropped foot stimulator. *Artificial Organs,* 21(3): 254-260.

Burschka, J. M., Keune, P.M., Menge, U., Hofstadt-van Oy, U., Oschmann, P., Hoos, O., (2012). An exploration of impaired walking dynamics and fatigue in multiple sclerosis. *BMC Neurology,* 12: 161.

Cameron, M. H. and Wagner, J. M., (2011). Gait abnormalities in multiple sclerosis: pathogenesis, evaluation, and advances in treatment. *Current Neurology and Neuroscience Reports,* 11(5): 507-515.

Caty, G. D., Detrembleur, C., Bleyenheuft, C., Lejeune, T.M., (2009). Reliability of lower limb kinematics, mechanics and energetics during gait in patients after stroke. *Journal of Rehabilitation Medicine,* 41(7): 588-590.

Cavanaugh, J. T., Gappmaier, V.O., Dibble, L.E., Gappmaier, E., (2011). Ambulatory activity in individuals with multiple sclerosis. *Journal of Neurologic Physical Therapy,* 35(1): 26-33.

Chaudhuri, A. and Behan, P. O., (2004). Fatigue in neurological disorders. *Lancet,* 363(9413): 978-988.

Chiaravalloti, N. D. and DeLuca, J., (2008). Cognitive impairment in multiple sclerosis. *Lancet Neurology,* 7(12): 1139-1151.

Cicchetti, D. V., (1994). Multiple comparison methods: establishing guidelines for their valid application in neuropsychological research. *Journal of Clinical and Experimental Neuropsycholy,* 16(1): 155-161.

Cleland, B. T., Ingraham, B.A., Pitluck, M.C., Woo, D., Ng, A.V., (2016). Reliability and Validity of Ratings of Perceived Exertion in Persons With Multiple Sclerosis. *Archives of Physical Medicine and Rehabilitation,* 97(6): 974-982.

Cofre Lizama, L.E., Khan, F., Lee, P.V., Galea, M.P., (2016). The use of laboratory gait analysis for understanding gait deterioration in people with multiple sclerosis. *Multiple Sclerosis,* 22(14): 1768-1776.

Coghe, G., Pilloni, G., Zucca, E., Porta, M., Corona, F., Frau, J., (2018). Exploring cognitive motor interference in multiple sclerosis by the visual Stroop test. *Multiple Sclerosis and Related Disorders,* 22: 8-11.

Cohen, A. B., Neema, M., Arora, A., Dell'oglio, E., Benedict, R.H., Tauhid, S., (2012). The relationships among MRI-defined spinal cord involvement, brain involvement, and disability in multiple sclerosis. *Journal of Neuroimaging,* 22(2): 122-128.

Cohen, J.A., Cutter, G.R., Fischer, J.S., Goodman, A.D., Heidenreich, F.R., Jak, A.J., et al., (2001). Use of the multiple sclerosis functional composite as an outcome measure in a phase 3 clinical trial. *Archives of Neurology,* 58(6): 961-967.

Cohen, J.A., Fischer, J.S., Bolibrush, D.M., Jak, A.J., Kniker, J.E., Mertz, L.A., et al., (2000). Intrarater and interrater reliability of the MS functional composite outcome measure. *Neurology,* 54(4): 802-806.

Cohen, J.A., Reingold, S.C., Polman, C.H., Wolinsky, J.S., (2012). Disability outcome measures in multiple sclerosis clinical trials: current status and future prospects. *Lancet Neurology,* 11(5): 467-476.

Coleman, C.I., Sobieraj, D.M., Marinucci, L.N., (2012). Minimally important clinical difference of the Timed 25-Foot Walk Test: results from a randomized controlled trial in patients with multiple sclerosis. *Current Medical Research and Opinion,* 28(1): 49-56.

Comber, L., Galvin, R., Coote, S., (2017). Gait deficits in people with multiple sclerosis: A systematic review and meta-analysis. *Gait & Posture,* 51: 25-35.

Compston, A. and Coles, A., (2008). Multiple sclerosis. *Lancet,* 372(9648): 1502-1517.

Correa, K.P., Devetak, G.F., Martello, S.K., de Almeida, J.C., Pauleto, A.C., Manffra, E.F., (2017). Reliability and Minimum Detectable Change of the Gait Deviation Index (GDI) in post-stroke patients. *Gait & Posture,* 53: 29-34.

Crayton, H.J. and Rossman, H.S., (2006). Managing the symptoms of multiple sclerosis: a multimodal approach. *Clinical Therapeutics,* 28(4): 445-460.

Crenshaw, S.J., Royer, T.D., Richards, J.G., Hudson, D.J., (2006). Gait variability in people with multiple sclerosis. *Multiple Sclerosis,* 12(5): 613-619.

Cutlip, R.G., Mancinelli, C., Huber, F., DiPasquale, J., (2000). Evaluation of an instrumented walkway for measurement of the kinematic parameters of gait. *Gait & Posture,* 12(2): 134-138.

Cutter, G. R., Baier, M.L., Rudick, R.A., Cookfair, D.L., Fischer, J.S., Petkau, J., et al., (1999). Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain,* 122 (Pt 5): 871-882.

Dahele, M., Skipworth, R.J., Wall, L., Voss, A., Preston, T., Fearon, K.C., (2007). Objective physical activity and self-reported quality of life in patients receiving palliative chemotherapy. *Journal of Pain and Symptom Management,* 33(6): 676-685.

Dalgas, U., Khølhede, T., Gijbels, D., Romberg, A., Santoyo, C., de Noordhout, B.M., et al., (2014). Aerobic intensity and pacing pattern during the six-minute walk test in patients with multiple sclerosis. *Journal of Rehabilitation Medecine,* 46(1): 59-66.

de Haan, A., de Ruiter, C.J., van Der Woude, L.H., Jongen, P.J., (2000). Contractile properties and fatigue of quadriceps muscles in multiple sclerosis. *Muscle & Nerve,* 23(10): 1534-1541.

de Vet, H.C., Terwee, C.B., Ostelo, R.W., Beckerman, H., Knol, D.L., Bouter, L.M., (2006a). Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health and Quality of Life Outcomes,* 4: 54.

de Vet, H.C., Terwee, C.B., Knol, D.L., Bouter, L.M., (2006b). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology,* 59(10): 1033-1039.

Delussu, A.S., Morone, G., Iosa, M., Bragoni, M., Paolucci, S., Traballesi, M., (2014). Concurrent validity of Physiological Cost Index in walking over ground and during robotic training in subacute stroke patients. *Biomed Research International,* 2014: 384896. doi: 10.1155/2014/384896.

Devetak, G.F., Martello, S.K., de Almeida, J.C., Correa, K.P., Iucksch, D.D., Manffra, E.F., (2016). Reliability and minimum detectable change of the gait profile score for post-stroke patients. *Gait & Posture,* 49: 382-387.

Dittner, A.J., Wessely, S.C., Brown, R.G., (2004). The assessment of fatigue: a practical guide for clinicians and researchers. *Journal of Psychosomatic Research,* 56(2): 157-170.

Dowd, K.P., Harrington, D.M., Donnelly, A.E., (2012). Criterion and concurrent validity of the activPAL professional physical activity monitor in adolescent females. *PLoS One,* 7(10): e47633.

Downing, A., Van Ryn, D., Fecko, A., Aiken, C., McGowan, S., Sawers, S., (2014). Effect of a 2-week trial of functional electrical stimulation on gait function and quality of life in people with multiple sclerosis. *International Journal of MS Care,* 16(3): 146-152.

Ehde, D.M., Gibbons, L.E., Chwastiak, L., Bombardier, C.H., Sullivan, M.D., Kraft, G.H., (2003). Chronic pain in a large community sample of persons with multiple sclerosis. *Multiple Sclerosis,* 9(6): 605-611.

Elbers, R.G., Rietberg, M.B., van Wegen, E.E., Verhoef, J., Kramer, S.F., Terwee, C.B., et al. (2012). Self-report fatigue questionnaires in multiple sclerosis, Parkinson's disease and stroke: a systematic review of measurement properties. *Quality of Life Research,* 21(6): 925-944.

Engelhard, M.M., Schmidt, K.M., Engel, C.E., Brenton, J.N., Patek, S.D., Goldman, M.D., (2016). The e-MSWS-12: improving the multiple sclerosis walking scale using item response theory. *Quality of Life Research,* 25(12): 3221-3230.

Esnouf, J.E., Taylor, P.N., Mann, G.E., Barrett, C.L., (2010). Impact on activities of daily living using a functional electrical stimulation device to improve dropped foot in people with multiple sclerosis, measured by the Canadian Occupational Performance Measure. *Multiple Sclerosis,* 16(9): 1141-1147.

Etemadi, Y., (2017). Dual task cost of cognition is related to fall risk in patients with multiple sclerosis: a prospective study. *Clinical Rehabilitation,* 31(2): 278-284.

Everaert, D.G., Thompson, A.K., Chong, S.L., Stein, R.B., (2010). Does functional electrical stimulation for foot drop strengthen corticospinal connections? *Neurorehabilitation Neural Repair,* 24(2): 168-177.

Faul, F., Erdfelder, E., Lang, A.G., Buchner, A., (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods,* 39: 175-191.

Fernandes, R., Armada-da-Silva, P., Pool-Goudzwaard, A.L., Moniz-Pereira, V., Veloso, A.P., (2016). Three dimensional multi-segmental trunk kinematics and kinetics during gait: Test-retest reliability and minimal detectable change. *Gait & Posture,* 46: 18-25.

Feys, P., Bibby, B., Romberg, A., Santoyo, C., Gebara, B., de Noordhout, B.M., et al., (2014). Within-day variability on short and long walking tests in persons with multiple sclerosis. *Journal of Neurological Sciences,* 338(1-2): 183-187.

Feys, P., Lamers, I., Francis, G., Benedict, R., Phillips, G., LaRocca,N., et al., (2017). The Nine-Hole Peg Test as a manual dexterity performance measure for multiple sclerosis. *Multiple Sclerosis,* 23(5): 711-720.

Feys, P., Severijns, D., Vantenderloo, S., Knuts, K., Hannes, D., Gijbels, D., et al., (2013). Spatio-temporal gait parameters change differently according to speed instructions and walking history in MS patients with different ambulatory dysfunction. *Multiple Sclerosis and Related Disorders,* 2(3): 238-246.

Filipovic Grcic, P., Matijaca, M., Lusic, I., Capkun, V., (2011). Responsiveness of walking-based outcome measures after multiple sclerosis relapses following steroid pulses. *Medical Science Monitor,* 17(12): Cr704-710.

Fischer, J.S., Rudick, R.A., Cutter, G.R., Reingold, S.C., (1999). The Multiple Sclerosis Functional Composite Measure (MSFC): an integrated approach to MS clinical outcome assessment. National MS Society Clinical Outcomes Assessmen Task Force. *Multiple Sclerosis,* 5(4): 244-250.

Flachenecker, P., Kumpfel, T., Kallmann, B., Gottschalk, M., Grauer, O., Rieckmann, P., et al., (2002). Fatigue in multiple sclerosis: a comparison of different rating scales and correlation to clinical parameters. *Multiple Sclerosis,* 8(6): 523-526.

Flegel, M., Knox, K., Nickel, D., (2012). Step-length variability in minimally disabled women with multiple sclerosis or clinically isolated syndrome. *International Journal of MS Care,* 14(1): 26-30.

Flensner, G., Ek, A.C., Soderhamn, O., Landtblom, A.M., (2011). Sensitivity to heat in MS patients: a factor strongly influencing symptomology--an explorative survey. *BMC Neurology,* 11: 27.

Franzen, M.D., Tishelman, A.C., Sharp, B.H., Friedman, A.G., (1987). An investigation of the test-retest reliability of the Stroop Color-Word Test across two intervals. *Archives of Clinical Neuropsychology,* 2(3): 265-272.

Fraser, C.L., Davagnanam, I., Radom, M., Plant, G.T., (2012). The time course and phenotype of Uhthoff phenomenon following optic neuritis. *Multiple Sclerosis,* 18(7): 1042-1044.

Freeman, J., Walters, R., Ingram, W., Slade, A., Hobart, J., Zajicek, J., (2013). Evaluating change in mobility in people with multiple sclerosis: relative responsiveness of four clinical measures. *Multiple Sclerosis,* 19(12): 1632-1639.

Friedman, P.J., (1990). Gait recovery after hemiplegic stroke. *International Disability Studies,* 12(3): 119-122.

Frohman, T.C., Davis, S.L., Beh, S., Greenberg, B.M., Remington, G., Frohman, E.M., (2013). Uhthoff's phenomena in MS--clinical features and pathophysiology. *Nature Reviews Neurology,* 9(9): 535-540.

Galea, M.P., Cofre Lizama, L.E., Butzkueven, H., Kilpatrick, T.J., (2017). Gait and balance deterioration over a 12-month period in multiple sclerosis patients with EDSS scores </= 3.0. *NeuroRehabilitation,* 40(2): 277-284.

Gianfrancesco, M.A., Triche, E.W., Fawcett, J.A., Labas, M.P., Patterson, T.S., Lo, A.C., (2011). Speed- and cane-related alterations in gait parameters in individuals with multiple sclerosis. *Gait & Posture,* 33(1): 140-142.

Gibson, J. and Frank, A., (2002). Supporting individuals with disabling multiple sclerosis. *Journal of the Royal Society of Medicine,* 95(12): 580-586.

Gijbels, D., Alders, G., Van Hoof, E., Charlier, C., Roelants, M., Broekmans, T., et al., (2010). Predicting habitual walking performance in multiple sclerosis: relevance of capacity and self-report measures. *Multiple Sclerosis,* 16(5): 618-626.

Gijbels, D., Dalgas, U., Romberg, A., de Groot, V., Bethoux, F., Vaney, C., et al., (2012). Which walking capacity tests to use in multiple sclerosis? A multicentre study providing the basis for a core set. *Multiple Sclerosis,* 18(3): 364-371.

Givon, U., Zeilig, G., Achiron, A., (2009). Gait analysis in multiple sclerosis: characterization of temporal-spatial parameters using GAITRite functional ambulation system. *Gait & Posture,* 29(1): 138-142.

Goldman, M.D., Marie, R.A., Cohen, J.A., (2008). Evaluation of the six-minute walk in multiple sclerosis subjects and healthy controls. *Multiple Sclerosis,* 14(3): 383-390.

Goldman, M.D., Marrie, R.A., Cohen, J.A., (2013). Clinically meaningful performance benchmarks in MS: timed 25-foot walk and the real world. *Neurology,* 81(21): 1856-1863.

Goodkin, D.E., Cookfair, D., Wende, K., Bourdette, D., Pullicino, P., Scherokman, B., et al., (1992). Inter- and intrarater scoring agreement using grades 1.0 to 3.5 of the Kurtzke Expanded Disability Status Scale (EDSS). Multiple Sclerosis Collaborative Research Group. *Neurology,* 42(4): 859-863.

Gosney, J.L., Scott, J.A., Snook, E.M., Motl, R.W., (2007). Physical activity and multiple sclerosis: validity of self-report and objective measures. *Family & Community Health,* 30(2): 144-150.

Gunn, H., Creanor, S., Haas, B., Marsden, J., Freeman, J., (2014). Frequency, characteristics, and consequences of falls in multiple sclerosis: findings from a cohort study. *Archives of Physical Medicine and Rehabilitation,* 95(3): 538-545.

Guyatt, G.H., Deyo, R.A., Charlson, M., Levine, M.N., Mitchell, A., (1989). Responsiveness and validity in health status measurement: a clarification. *Journal of Clinical Epidemiology,* 42(5): 403-408.

Guyatt, G.H., Sackett, D.L., Cook, D.J., (1993). Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *Jama,* 270(21): 2598-2601.

Hamacher, D., Herold, F., Wiegel, P., Hamacher, D., Schega, L., (2015). Brain activity during walking: A systematic review. *Neuroscience & Biobehavioral Reviews,* 57: 310-327.

Hamalainen, P., Ikonen, A., Romberg, A., Helenius, H., Ruutiainen, J., (2012). The effects of heat stress on cognition in persons with multiple sclerosis. *Multiple Sclerosis,* 18(4): 489-497.

Hamilton, F., Rochester, L., Paul, L., Rafferty, D., O'Leary, C.P., Evans, J.J., (2009). Walking and talking: an investigation of cognitive-motor dual tasking in multiple sclerosis. *Multiple Sclerosis,* 15(10): 1215-1227.

Harrington, D.M., Dwod, K.P., Tudor-Locke, C., Donnelly, A.E., (2012). A steps/minute value for moderate intensity physical activity in adolescent females. *Pediatric Exercise Science,* 24(3): 399-408.

Hauser, S.L. and Oksenberg, J.R., (2006). The neurobiology of multiple sclerosis: genes, inflammation, and neurodegeneration. *Neuron,* 52(1): 61-76.

Heesen, C., Bohm, J., Reich, C., Kasper, J., Goebel, M., Gold, S.M., (2008). Patient perception of bodily functions in multiple sclerosis: gait and visual function are the most valuable. *Multiple Sclerosis,* 14(7): 988-991.

Heine, M., van den Akker, L.E., Verschuren, O., Visser-Meily, A., Kwakkel, G., (2015). Reliability and responsiveness of cardiopulmonary exercise testing in fatigued persons with multiple sclerosis and low to mild disability. *PLoS One,* 10(3): e0122260.

Hobart, J., Blight, A.R., Goodman, A., Lynn, F., Putzki, N., (2013). Timed 25-foot walk: direct evidence that improving 20% or greater is clinically meaningful in MS. *Neurology,* 80(16): 1509-1517.

Hobart, J., Freeman, J., Thompson, A., (2000). Kurtzke scales revisited: the application of psychometric methods to clinical intuition. *Brain,* 123 (Pt 5): 1027-1040.

Hobart, J., Kalkers, N., Barkhof, F., Uitdehaag, B., Polman, C., Thompson, A., (2004). Outcome measures for multiple sclerosis clinical trials: relative measurement precision of the Expanded Disability Status Scale and Multiple Sclerosis Functional Composite. *Multiple Sclerosis,* 10(1): 41-46.

Hobart, J.C., Riazi, A., Lamping, D.L., Fitzpatrick, R., Thompson, A.J., (2003). Measuring the impact of MS on walking ability: the 12-Item MS Walking Scale (MSWS-12). *Neurology,* 60(1): 31-36.

Hoogervorst, E.L., Kalkers, N., Uitdehaag, B.M., Polman, C.H., (2002). A study validating changes in the multiple sclerosis functional composite. *Archives of Neurology,* 59(1): 113-116.

Hopkins, W.G., (2000). Measures of reliability in sports medicine and science. *Sports Medicine,* 30(1): 1-15.

Huisinga, J.M., Schmid, K.K., Filipi, M.L., Stergiou, N., (2013). Gait mechanics are different between healthy controls and patients with multiple sclerosis. *Journal of Applied Biomechanics,* 29(3): 303-311.

Iriarte, J. and de Castro, P., (1998). Correlation between symptom fatigue and muscular fatigue in multiple sclerosis. *European Journal of Neurology,* 5(6): 579-585.

Jensen, H.B., Mamoei, S., Ravnborg, M., Dalgas, U., Stenager, E., (2016). Distribution-based estimates of minimum clinically important difference in cognition, arm function and lower body function after slow release-fampridine treatment of patients with multiple sclerosis. *Multiple Sclerosis and Related Disorders,* 7: 58-60.

Jordan, B., Mehl, T., Schweden, T.L.K., Menge, U., Zierz, S., (2017). Assessment of physical fatigability and fatigue perception in myasthenia gravis. *Muscle & Nerve,* 55(5): 657-663.

Kadaba, M.P., Ramakrishnan, H.K., Wooten, M.E., (1990). Measurement of lower extremity kinematics during level walking. *Journal of Orthopaedic Research,* 8(3): 383-392.

Kadaba, M.P., Ramakrishnan, H.K., Wooten, M.E., Gainey, J., Gorton, G., Cochran, G.V., (1989). Repeatability of kinematic, kinetic, and electromyographic data in normal adult gait. *Journal of Orthopaedic Research,* 7(6): 849-860.

Kafri, M. and Laufer, Y., (2015). Therapeutic effects of functional electrical stimulation on gait in individuals post-stroke. *Annals of Biomedical Engineering,* 43(2): 451-466.

Kainz, H., Carty, C.P., Maine, S., Walsh, H.P.J., Lloyd, D.G., Modenese, L., (2017a). Effects of hip joint centre mislocation on gait kinematics of children with cerebral palsy calculated using patient-specific direct and inverse kinematic models. *Gait & Posture,* 57: 154-160.

Kainz, H., Graham, D., Edwards, J., Walsh, H.P.J., Maine, S., Boyd, R.N., et al., (2017b). Reliability of four models for clinical gait analysis. *Gait & Posture,* 54: 325-331.

Kaipust, J.P., Huisinga, J.M., Filipi, M., Stergiou, N., (2012). Gait variability measures reveal differences between multiple sclerosis patients and healthy controls. *Motor Control*, 16(2): 229-244.

Kalkers, N.F., Bergers, L., de Groot, V., Lazeron, R.H., van Walderveen, M.A., Uitdehaag, B.M., et al., (2001). Concurrent validity of the MS Functional Composite using MRI as a biological disease marker. *Neurology*, 56(2): 215-219.

Kalmar, J.H., Gaudino, E.A., Moore, N.B., Halper, J., Deluca, J., (2008). The relationship between cognitive deficits and everyday functional activities in multiple sclerosis. *Neuropsychology*, 22(4): 442-449.

Kalron, A., (2016). Gait variability across the disability spectrum in people with multiple sclerosis. *Journal of Neurological Sciences*, 361: 1-6.

Kalron, A., Achiron, A., Dvir, Z., (2011). Muscular and gait abnormalities in persons with early onset multiple sclerosis. *Journal of Neurologic Physical Therapy*, 35(4): 164-169.

Kalron, A., Dvir, Z., Achiron, A., (2010). Walking while talking--difficulties incurred during the initial stages of multiple sclerosis disease process. *Gait & Posture*, 32(3): 332-335.

Kalron, A., Dvir, Z., Frid, L., Achiron, A., (2013). Quantifying gait impairment using an instrumented treadmill in people with multiple sclerosis. *ISRN Neurology*, 2013: 867575.

Kalron, A., Dvir, Z., Givon, U., Baransi, H., Achiron, A., (2014). Gait and jogging parameters in people with minimally impaired multiple sclerosis. *Gait & Posture*, 39(1): 297-302.

Kalron, A., Frid, L., Menascu, S., (2017). Gait Characteristics in Adolescents With Multiple Sclerosis. *Pediatric Neurology*, 68: 73-76.

Kasser, S.L., Goldstein, A., Wood, P.K., Sibold, J., (2017). Symptom variability, affect and physical activity in ambulatory persons with multiple sclerosis: Understanding patterns and time-bound relationships. *Disability and Health Journal*, 10(2): 207-213.

Kaufman, K., Miller, E., Kingsbury, T., Russell Esposito, E., Wolf, E., Wilken, J., et al., (2016). Reliability of 3D gait data across multiple laboratories. *Gait & Posture,* 49: 375-381.

Kaufman, M., Moyer, D., Norton, J., (2000). The significant change for the Timed 25-foot Walk in the multiple sclerosis functional composite. *Multiple Sclerosis,* 6(4): 286-290.

Kelleher, K.J., Spence, W., Solomonidis, S., Apatsidis, D., (2010). The characterisation of gait patterns of people with multiple sclerosis. *Disability Rehabilitation,* 32(15): 1242-1250.

Kempen, J.C., de Groot, V., Knol, D.L., Polman, C.H., Lankhorst, G.J., Beckerman, H., (2011). Community walking can be assessed using a 10-metre timed walk test. *Multiple Sclerosis,* 17(8): 980-990.

Kempen, J.C., Doorenbosch, C.A., Knol, D.L., de Groot, V., Beckerman, H., (2016). Newly Identified Gait Patterns in Patients With Multiple Sclerosis May Be Related to Push-off Quality. *Physical Therapy,* 96(11): 1744-1752.

Kent-Braun, J.A., Sharma, K.R., Weiner, M.W., Miller, R.G., (1994). Effects of exercise on muscle activation and metabolism in multiple sclerosis. *Muscle & Nerve,* 17(10): 1162-1169.

Kesar, T.M., Binder-Macleod, S.A., Hicks, G.E., Reisman, D.S., (2011). Minimal detectable change for gait variables collected during treadmill walking in individuals post-stroke. *Gait & Posture,* 33(2): 314-317.

Kierkegaard, M., Einarsson, U., Gottberg, K., von Koch, L., Holmqvist, L.W., (2012). The relationship between walking, manual dexterity, cognition and activity/participation in persons with multiple sclerosis. *Multiple Sclerosis,* 18(5): 639-646.

Kieseier, B.C. and Pozzilli, C., (2012). Assessing walking disability in multiple sclerosis. *Multiple Sclerosis,* 18(7): 914-924.

Klejman, S., Andrysek, J., Dupuis, A., Wright, V., (2010). Test-retest reliability of discrete gait parameters in children with cerebral palsy. *Archives of Physical Medicine and Rehabilitation,* 91(5): 781-787.

Kluger, B.M., Krupp, L.B., Enoka, R.M., (2013). Fatigue and fatigability in neurologic illnesses: proposal for a unified taxonomy. *Neurology,* 80(4): 409-416.

Koo, T.K. and Li, M.Y., (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine,* 15(2): 155-163.

Kos, D., Kerckhofs, E., Ketelaer, P., Duportail, M., Nagels, G., D'Hooghe, M., et al., (2004). Self-report assessment of fatigue in multiple sclerosis: a critical evaluation. *Occupational Therapy in Health Care,* 17(3-4): 45-62.

Kos, D., Raeymaekers, J., Van Remoorterl, A., D'Hooghe, M., Nagels, G., D'Haeseleer, M., et al., (2017). Electronic visual analogue scales for pain, fatigue, anxiety and quality of life in people with multiple sclerosis using smartphone and tablet: a reliability and feasibility study. *Clinical Rehabilitation,* 31(9): 1215-1225.

Kragt, J.J., Thompson, A.J., Montalban, X., Tintore, M., Rio, J., Polman, C.H., et al., (2008). Responsiveness and predictive value of EDSS and MSFC in primary progressive MS. *Neurology,* 70(13 Pt 2): 1084-1091.

Kragt, J.J., van der Linden, F.A., Nielsen, J.M., Uitdehaag, B.M., Polman, C.H., (2006). Clinical impact of 20% worsening on Timed 25-foot Walk and 9-hole Peg Test in multiple sclerosis. *Multiple Sclerosis,* 12(5): 594-598.

Krupp, L.B., LaRocca, N.G., Muir-Nash, J., Steinberg, A.D., (1989). The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology,* 46(10): 1121-1123.

Kurtzke, J.F., (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology,* 33(11): 1444-1452.

Langeskov-Christensen, M., Bisson, E.J., Finlayson, M.L., Dalgas, U., (2017). Potential pathophysiological pathways that can explain the positive effects of exercise on fatigue

in multiple sclerosis: A scoping review. *Journal of the Neurological Sciences,* 373: 307-320.

Larocca, N.G., (2011). Impact of walking impairment in multiple sclerosis: perspectives of patients and care partners. *Patient,* 4(3): 189-201.

Laroche, D., Duval, A., Morisset, C., Beis, J.N., d'Athis, P., Maillefert, J.F., et al., (2011). Test-retest reliability of 3D kinematic gait variables in hip osteoarthritis patients. *Osteoarthritis Cartilage,* 19(2): 194-199.

Larson, R.D., Larson, D.J., Baumgartner, T.B., White, L.J., (2013). Repeatability of the timed 25-foot walk test for individuals with multiple sclerosis. *Clinical Rehabilitation,* 27(8): 719-723.

Lassmann, H., (2012). Cortical lesions in multiple sclerosis: inflammation versus neurodegeneration. *Brain,* 135(Pt 10): 2904-2905.

Learmonth, Y.C., Dlugonski, D.D., Pilutti, L.A., Sandroff, B.M., Motl, R.W., (2013a). The reliability, precision and clinically meaningful change of walking assessments in multiple sclerosis. *Multiple Sclerosis,* 19(13): 1784-1791.

Learmonth, Y.C., Dlugonski, D.D., Pilutti, L.A., Sandroff, B.M., Klaren, R., Motl, R.W., (2013b). Psychometric properties of the Fatigue Severity Scale and the Modified Fatigue Impact Scale. *Journal of the Neurological Sciences,* 331(1-2): 102-107.

Learmonth, Y.C., Ensari, I., Motl, R.W., (2017). Cognitive Motor Interference in Multiple Sclerosis: Insights From a Systematic Quantitative Review. *Archives of Physical Medicine and Rehabilitation,* 98(6): 1229-1240.

Learmonth, Y.C., Paul, L., McFadyen, A.K., Mattison, P., Miller, L., (2012). Reliability and clinical significance of mobility and balance assessments in multiple sclerosis. *International Journal of Rehabilitation Research,* **35**(1): 69-74.

Learmonth, Y. C., Sandroff, B.M., Pilutti, L.A., Klaren, R.E., Ensari, I., Riskin, B.J., et al., (2014). Cognitive motor interference during walking in multiple sclerosis using an alternate-letter alphabet task. *Archives of Physical Medicine and Rehabilitation,* 95(8): 1498-1503.

Lee, K.A., Hicks, G., Nino-Murcia, G., (1991). Validity and reliability of a scale to assess fatigue. *Psychiatry Research,* 36(3): 291-298.

Leon, G.R., Koscheyev, V.S., Stone, E.A., (2008). Visual analog scales for assessment of thermal perception in different environments. *Aviation, Space, and Environmental Medicine,* 79(8): 784-786.

Leone, C., Severijns, D., Dolezalova, V., Baert, I., Dalgas, U., Romberg, A., et al., (2016). Prevalence of Walking-Related Motor Fatigue in Persons With Multiple Sclerosis: Decline in Walking Distance Induced by the 6-Minute Walk Test. *Neurorehabilitation Neural Repair,* 30(4): 373-383.

Lindemann, U., Klenk, J., Becker, C., (2016). Assessment of fatigability of older women during sit-to-stand performance. *Aging Clinical and Experimental Research,* 28(5): 889-893.

Liu, Y., Zhu, S.H., Wang, G.H., Ye, F., Li, P.Z., (2013). Validity and reliability of multiparameter physiological measurements recorded by the Equivital LifeMonitor during activities of various intensities. *Journal of Occupational and Environmental Hygiene,* 10(2): 78-85.

Lord, S.E., Halligan, P.W., Wade, D.T., (1998). Visual gait analysis: the development of a clinical assessment and scale. *Clinical Rehabilitation,* 12(2): 107-119.

Lublin, F.D. and Reingold, S.C., (1996). Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. *Neurology,* 46(4): 907-911.

Lublin, F.D., Reingold, S.C., Cohen, J.A., Cutter, G.R., Sørensen, P.S., Thompson, A.J., et al., (2014). Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology,* 83(3): 278-286.

Lukas, C., Sombekke, M.H., Bellenberg, B., Hahn, H.K., Popescu, V., Bendfeldt, K., et al., (2013). Relevance of spinal cord abnormalities to clinical disability in multiple sclerosis: MR imaging findings in a large cohort of patients. *Radiology,* 269(2): 542-552.

Mackenzie, I.S., Morant, S.V., Bloomfield, G.A., MacDonald, T.M., O'Riordan, J., (2014). Incidence and prevalence of multiple sclerosis in the UK 1990-2010: a descriptive study in the General Practice Research Database. *Journal of Neurology, Neurosurgery, and Psychiatry,* 85(1): 76-84.

Mackey, A.H., Walt, S.E., Lobb, G.A., Stott, N.S., (2005). Reliability of upper and lower limb three-dimensional kinematics in children with hemiplegia. *Gait & Posture,* 22(1): 1-9.

MacLeod, C.M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin,* 109(2): 163-203.

Marrie, R.A., Cutter, G., Tyry, T., Vollmer, T., Campagnolo, D., (2007). Disparities in the management of multiple sclerosis-related bladder symptoms. *Neurology,* 68(23): 1971-1978.

Martin, C.L., Phillips, B.A., Kilpatrick, T.J., Butzkueven, H., Tubridy, N., McDonald, E., et al., (2006). Gait and balance impairment in early multiple sclerosis in the absence of clinical disability. *Multiple Sclerosis,* 12(5): 620-628.

Mayer, L., Warring, T., Agrella, S., Rogers, H.L., Fox, E.J., (2015). Effects of functional electrical stimulation on gait function and quality of life for people with multiple sclerosis taking dalfampridine. *International Journal of MS Care,* 17(1): 35-41.

Maynard, V., Bakheit, A.M., Oldham, J., Freeman, J., (2003). Intra-rater and inter-rater reliability of gait measurements with CODA mpx30 motion analysis system. *Gait & Posture,* 17(1): 59-67.

McDonald, W.I. and Sears, T.A., (1970). The effects of experimental demyelination on conduction in the central nervous system. *Brain,* 93(3): 583-598.

McGinley, J.L., Baker, R., Wolfe, R., Morris, M.E., (2009). The reliability of three-dimensional kinematic gait measurements: a systematic review. *Gait & Posture,* 29(3): 360-369.

McGuigan, C. and Hutchinson, M., (2004). Confirming the validity and responsiveness of the Multiple Sclerosis Walking Scale-12 (MSWS-12). *Neurology,* 62(11): 2103-2105.

McLoughlin, J.V., Barr, C.J., Patritti, B., Crotty, M., Lord, S.R., Sturnieks, D.L., (2016). Fatigue induced changes to kinematic and kinetic gait parameters following six minutes of walking in people with multiple sclerosis. *Disability Rehabilitation,* 38(6): 535-543.

McLoughlin, J.V., Lord, S.R., Barr, C.J., Crotty, M., Sturnieks, D.L., (2015). Dorsiflexion assist orthosis reduces the physiological cost and mitigates deterioration in strength and balance associated with walking in people with multiple sclerosis. *Archives of Physical Medicine and Rehabilitation,* 96(2): 226-232.e221.

Meldrum, D., Shouldice, C., Conroy, R., Jones, K., Forward, M., (2014). Test-retest reliability of three dimensional gait analysis: including a novel approach to visualising agreement of gait cycle waveforms with Bland and Altman plots. *Gait & Posture,* 39(1): 265-271.

Miller, D., Barkhof, F., Montalban, X., Thompson, A., Filipi, M., (2005). Clinically isolated syndromes suggestive of multiple sclerosis, part I: natural history, pathogenesis, diagnosis, and prognosis. *Lancet Neurology,* 4(5): 281-288.

Miller, D.M., Rudick, R.A., Cutter, G., Baier, M., Fischer, J.S., (2000). Clinical significance of the multiple sclerosis functional composite: relationship to patient-reported quality of life. *Archives of Neurology,* 57(9): 1319-1324.

Miller, L., McFadyen, A., Lord, A.C., Hunter, R., Paul, L., Rafferty, D., et al., (2017). Functional Electrical Stimulation for Foot Drop in Multiple Sclerosis: A Systematic Review and Meta-Analysis of the Effect on Gait Speed. *Archives of Physical Medicine and Rehabilitation,* 98(7): 1435-1452.

Miller, L., Rafferty, D., Paul, L., Mattison, P., (2016). The impact of walking speed on the effects of functional electrical stimulation for foot drop in people with multiple sclerosis. *Disability Rehabilitation: Assistive Technology,* 11(6): 478-483.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International Journal of Surgery,* 8(5): 336-341.

Mohr, D.C., Goodkin, D.E., Gatto, N., Van der Wende, J., (1997). Depression, coping and level of neurological impairment in multiple sclerosis. *Multiple Sclerosis,* 3(4): 254-258.

Mokkink, L.B., Galindo-Garre, F., Uitdehaag, B.M., (2016). Evaluation of the Multiple Sclerosis Walking Scale-12 (MSWS-12) in a Dutch sample: Application of item response theory. *Multiple Sclerosis,* 22(14): 1867-1873.

Mokkink, L.B., Terwee, C.B., Gibbons, E., Stratford, P.W., Alonso, J., Patrick, D.L., et al., (2010a). Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. *BMC Medical Research Methodology*, 10: 82.

Mokkink, L.B., Terwee, B.C., Patrick, D.L., Alonso, J., Stratford, P.W., Knol,D.L., et al., (2010b). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology,* 63(7): 737-745.

Mokkink, L.B., Terwee, B.C., Patrick, D.L., Alonso, J., Stratford, P.W., Knol,D.L., et al., (2010c). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research,* 19(4): 539-549.

Mokkink, L.B., Terwee, C.B., Prinsen, C.A., de Vet, H.C., (2016). Taxonomy of measurement properties: A commentary on Polit (2015). *International Journal of Nursing Studies,* 53: 399-400.

Monaghan, K., Delahunt, E., Caulfield, B., (2007). Increasing the number of gait trial recordings maximises intra-rater reliability of the CODA motion analysis system. *Gait & Posture,* 25(2): 303-315.

Monticone, M., Ambrosini, E., Fiorentini, R., Rocca, B., Liquori, V., Pedrocchi, A., et al., (2014). Reliability of spatial-temporal gait parameters during dual-task interference in people with multiple sclerosis. A cross-sectional study. *Gait & Posture,* 40(4): 715-718.

Morris, M.E., Cantwell, C., Vowels, L., Dodd, K., (2002). Changes in gait and fatigue from morning to afternoon in people with multiple sclerosis. *Journal of Neurology, Neurosurgery, and Psychiatry,* 72(3): 361-365.

Morrissey, S.P., Miller, D.H., Kendall, B.E., Kingsley, D.P., Kelly, M.A., Francis, D.A., et al., (1993). The significance of brain magnetic resonance imaging abnormalities at presentation with clinically isolated syndromes suggestive of multiple sclerosis. A 5-year follow-up study. *Brain,* 116 (Pt 1): 135-146.

Motl, R.W., Dlugonski, D., Suh, Weikert, M., Agiovlasitis, S., Fernhall, B., et al., (2010). Multiple Sclerosis Walking Scale-12 and oxygen cost of walking. *Gait & Posture,* 31(4): 506-510.

Motl, R.W., Learmonth, Y.C., Pilutti, L.A., Dlugonski, D., Klaren, R., (2014). Validity of minimal clinically important difference values for the multiple sclerosis walking scale-12? *European Neurology,* 71(3-4): 196-202.

Motl, R.W., McAuley, E., Mullen, S., (2011). Longitudinal measurement invariance of the Multiple Sclerosis Walking Scale-12. *Journal of the Neurological Sciences,* 305(1-2): 75-79.

Motl, R.W., Sandroff, B.M., DeLuca, J., (2016). Exercise Training and Cognitive Rehabilitation: A Symbiotic Approach for Rehabilitating Walking and Cognitive Functions in Multiple Sclerosis? *Neurorehabilitation Neural Repair,* 30(6): 499-511.

Motl, R.W. and Snook, E.M., (2008). Confirmation and extension of the validity of the Multiple Sclerosis Walking Scale-12 (MSWS-12). *Journal of the Neurological Sciences,* 268(1-2): 69-73.

Motl, R.W., Sosnoff, J.J., Dlugonski, D., Pilutti, L.A., Klaren, R., Sandroff, B.M., (2014). Walking and cognition, but not symptoms, correlate with dual task cost of walking in multiple sclerosis. *Gait & Posture,* 39(3): 870-874.

Motl, R.W., Zhu, W., Park, Y., McAuley, E., Scott, J.A., Snook, E.M., (2007). Reliability of scores from physical activity monitors in adults with multiple sclerosis. A*dapted Physical Activity Quarterly,* 24(3): 245-253.

Mount, J. and Dacko, S., (2006). Effects of dorsiflexor endurance exercises on foot drop secondary to multiple sclerosis: a pilot study. *NeuroRehabilitation,* 21(1): 43-50.

Nauta, I.M., Balk, L.J., Sonder, J.M., Hulst, H.E., Uitdehaag, B.M., Fasotti, L., et al., (2018). The clinical value of the patient-reported multiple sclerosis neuropsychological

screening questionnaire. *Multiple Sclerosis,* 1352458518777295. doi:10.1177/1352458518777295.

Ng, A.V., Miller, R.G., Gelinas, D., Kent-Braun, J.A., (2004). Functional relationships of central and peripheral muscle alterations in multiple sclerosis. *Muscle & Nerve,* 29(6): 843-852.

Nieuwenhuys, A., Papageorgiou, E., Molenaers, G., Monari, D., de Laet, T., Desloovere, K., (2017). Inter- and intrarater clinician agreement on joint motion patterns during gait in children with cerebral palsy. *Developmental Medicine & Child Neurology,* 59(7): 750-755.

Nilsagard, Y., Lundholm, C., Gunnarsson, L.G., Dcnison, E., (2007). Clinical relevance using timed walk tests and 'timed up and go' testing in persons with multiple sclerosis. *Physiotherapy Research International,* 12(2): 105-114.

Nogueira, L.A., Teixeira, L., Sabino, P., Filho, H.A., Alvarenga, R.M., Santos Thuler, L.C., (2013). Gait characteristics of multiple sclerosis patients in the absence of clinical disability. *Disability Rehabilitation,* 35(17): 1472-1478.

Noonan, K.J., Halliday, S., Browne, R., O'Brien, S., Kayes, K., Feinberg, J., (2003). Interobserver variability of gait analysis in patients with cerebral palsy. *Journal of Pediatric Orthopaedics,* 23(3): 279-287; discussion 288-291.

Nortvedt, M.W., Riise, T., Myhr, K.M., Landtblom, A.M., Bakke, A., Nyland, H.I., (2001). Reduced quality of life among multiple sclerosis patients with sexual disturbance and bladder dysfunction. *Multiple Sclerosis,* 7(4): 231-235.

Oervik, M.S., Sejbaek, T., Penner, I.K., Roar, M., Blaabjerg,M., (2017). Validation of the fatigue scale for motor and cognitive functions in a danish multiple sclerosis cohort. *Multiple Sclerosis and Related Disorders,* 17: 130-134.

Paisley, S., Beard, S., Hunn, A., Wight, J., (2002). Clinical effectiveness of oral treatments for spasticity in multiple sclerosis: a systematic review. *Multiple Sclerosis,* 8(4): 319-329.

Paltamaa, J., Sarasoja, T., Leskinen, E., Wikström, J., Mälkiä, E., (2008). Measuring deterioration in international classification of functioning domains of people with multiple sclerosis who are ambulatory. *Physical Therapy,* 88(2): 176-190.

Paltamaa, J., West, H., Sarasoja, T., Wikström, J., Mälkiä, E., (2005). Reliability of physical functioning measures in ambulatory subjects with MS. *Physiotherapy Research International,* 10(2): 93-109.

Panicker, J.N. and Fowler, C.J., (2015). Lower urinary tract dysfunction in patients with multiple sclerosis. *Handbook of Clinical Neurology,* 130: 371-381.

Pashler, H. (1994). Graded capacity-sharing in dual-task interference? *Journal of Experimental Psychology: Human Perception and Performance,* 20(2): 330-342.

Patten, S.B., Burton, J.M., Fiest, K.M., Wiebe, S., Bulloch, A.G., Koch, M., et al., (2015). Validity of four screening scales for major depression in MS. *Multiple Sclerosis,* 21(8): 1064-1071.

Patti, F. (2009). Cognitive impairment in multiple sclerosis. *Multiple Sclerosis,* 15(1): 2-8.

Pau, M., Coghe, G., Atzeni, C., Corona, F., Pilloni, G., Marrosu, M.G., et al., (2014). Novel characterization of gait impairments in people with multiple sclerosis by means of the gait profile score. *Journal of the Neurological Sciences,* 345(1-2): 159-163.

Paul, L., Rafferty, D., Young, S., Miller, L., Mattison, P., McFadyen, A., (2008). The effect of functional electrical stimulation on the physiological cost of gait in people with multiple sclerosis. *Multiple Sclerosis,* 14(7): 954-961.

Penko, A.L., Barkley, J.E., Koop, M.M., Alberts, J.L., (2017). Borg scale is valid for ratings of perceived exertion for individuals with Parkinson's disease. *International Journal of Exercise Science,* 10(1): 76-86.

Penner, I.K., Raselli, C., Stöcklin, M., Opwis, K., Kappos, L., Calabrese, P., (2009). The Fatigue Scale for Motor and Cognitive Functions (FSMC): validation of a new instrument to assess multiple sclerosis-related fatigue. *Multiple Sclerosis,* 15(12): 1509-1517.

Peruzzi, A., Cereatti, A., Della Croce, U., Mirelman, A., (2016). Effects of a virtual reality and treadmill training on gait of subjects with multiple sclerosis: a pilot study. *Multiple Sclerosis and Related Disorders,* 5: 91-96.

Peterson, E.W., Cho, C.C., Finlayson, M.L., (2007). Fear of falling and associated activity curtailment among middle aged and older adults with multiple sclerosis. *Multiple Sclerosis,* 13(9): 1168-1175.

Phan-Ba, R., Calay, P., Grodent, P., Delcrue, G., Lommers, E., Delvaux, V., et al., (2012). Motor fatigue measurement by distance-induced slow down of walking speed in multiple sclerosis. *PLoS One,* 7(4): e34744.

Phan-Ba, R., Pace, A., Calay, P., Grodent, P., Douchamps, F., Hyde, R., et al., (2011). Comparison of the timed 25-foot and the 100-meter walk as performance measures in multiple sclerosis. *Neurorehabilitation Neural Repair,* 25(7): 672-679.

Pilutti, L.A., Dlugonski, D., Sandroff, B.M., Suh, Y., Pula, J.H., Sosnoff, J.J., et al., (2013). Further validation of multiple sclerosis walking scale-12 scores based on spatiotemporal gait parameters. *Archives of Physical Medicine and Rehabilitation,* 94(3): 575-578.

Pistacchi, M., Gioulis, M., Sanson, F., De Giovannini, E., Filippi, G., Rossetto, F., et al., (2017). Gait analysis and clinical correlations in early Parkinson's disease. *Functional Neurology,* 32(1): 28-34.

Polit, D.F. (2016). Taxonomy of measurement properties: A response to the commentary by Mokkink et al. (2015). *International Journal of Nursing Studies,* 53: 401-402.

Polman, C.H., Reingold, S.C., Banwell, B., Clanet, M., Cohen, J.A., Filippi, M., et al., (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of Neurology,* 69(2): 292-302.

Powell, D.J.H., Liossi, C., Schlotz, W., Moss-Morris, R., (2017). Tracking daily fatigue fluctuations in multiple sclerosis: ecological momentary assessment provides unique insights. *Journal of Behavioral Medicine,* 40(5): 772-783.

Prinsen, C.A.C., Mokkink, L.B., Bouter, L.M., Alonso, J., Patrick, D.L., de Vet, H.C.W., et al., (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research,* 27(5): 1147-1157.

Proessl, F., Ketelhut, N.B., Rudroff, T., (2018). No association of leg strength asymmetry with walking ability, fatigability, and fatigue in multiple sclerosis. *International Journal of Rehabilitation Research,* 41(3): 267-269.

Raja, K., Joseph, B., Benjamin, S., Minocha, V., Rana, B., (2007). Physiological cost index in cerebral palsy: its role in evaluating the efficiency of ambulation. *Journal of Pediatric Orthopaedics,* 27(2): 130-136.

Rao, S.M., Leo, G.J., Ellington, L., Nauertz, T., Bernardin, L., Unverzagt, F., (1991). Cognitive dysfunction in multiple sclerosis. II. Impact on employment and social functioning. *Neurology,* 41(5): 692-696.

Rasmussen, H.M., Nielsen, D.B., Pedersen, N.W., Overgaard, S., Holsgaard-Larsen, A., (2015). Gait Deviation Index, Gait Profile Score and Gait Variable Score in children with spastic cerebral palsy: Intra-rater reliability and agreement across two repeated sessions. *Gait & Posture,* 42(2): 133-137.

Redekop, S., Andrysek, J., Wright, V., (2008). Single-session reliability of discrete gait parameters in ambulatory children with cerebral palsy based on GMFCS level. *Gait & Posture,* 28(4): 627-633.

Remelius, J.G., Jones, S.L., House, J.D., Busa, M.A., Averill, J.L., Sugumaran, K., et al., (2012). Gait impairments in persons with multiple sclerosis across preferred and fixed walking speeds. *Archives of Physical Medicine and Rehabilitation,* 93(9): 1637-1642.

Renfrew, M.L., Flowers, P., Lord, A.C., Rafferty, D., McFadyen, A.K., Bowers, R., et al., (2018). An exploration of the experiences and utility of functional electrical stimulation for foot drop in people with multiple sclerosis. *Disability Rehabilitation,* 1-9. doi: 10.1080/09638288.2018.1501100.

Rietberg, M.B., Van Wegen, E.E., Kwakkel, G., (2010). Measuring fatigue in patients with multiple sclerosis: reproducibility, responsiveness and concurrent validity of three Dutch self-report questionnaires. *Disability Rehabilitation,* 32(22): 1870-1876.

Rizzo, M.A., Hadjimichael, O.C., Preiningerova, J., Vollmer, T.L., (2004). Prevalence and treatment of spasticity reported by multiple sclerosis patients. *Multiple Sclerosis,* 10(5): 589-595.

Roiz Rde, M., Cacho, E.W., Pazinatto, M.M., Reis, J.G., Cliquet, A. Jr., Barasnevicius-Quagliato, E.M., (2010). Gait analysis comparing Parkinson's disease with healthy elderly subjects. *Arquivos de Neuro-Psiquiatria,* 68(1): 81-86.

Romani, A., Bergamaschi, R., Candeloro, E., Alfonsi, E., Callieco, R., Cosi, V., (2004). Fatigue in multiple sclerosis: multidimensional assessment and response to symptomatic treatment. *Multiple Sclerosis,* 10(4): 462-468.

Rooney, S., Wood, L., Moffat, F., Paul, L., (2019). Is Fatigue Associated With Aerobic Capacity and Muscle Strength in People With Multiple Sclerosis: A Systematic Review and Meta-analysis. *Archives of Physical Medicine and Rehabilitation,* doi: 10.1016/j.apmr.2019.06.014.

Rosti-Otajarvi, E., Hamalainen, P., Koivisto, K., Hokkanen, L., (2008). The reliability of the MSFC and its components. *Acta Neurologica Scandinavica,* 117(6): 421-427.

Rovaris, M. and Filippi, M., (2000). MRI correlates of cognitive dysfunction in multiple sclerosis patients. *Journal of Neurovirology,* 6 Suppl 2: S172-175.

Rudroff, T., Kindred, J.H., Ketelhut, N.B., (2016). Fatigue in Multiple Sclerosis: Misconceptions and Future Research Directions. *Frontiers in Neurology,* 7: 122.

Sakakibara, R. (2019). Neurogenic lower urinary tract dysfunction in multiple sclerosis, neuromyelitis optica, and related disorders. *Clinical Autonomic Research,* 29(3): 313-320.

Sale, D., (2004). Postactivation potentiation: role in performance. *British Journal of Sports Medicine,* 38(4): 386-387.

Sandroff, B.M., Motl, R.W., Davis, S.L., (2016). Effects of vigorous walking exercise on core body temperature and inhibitory control in thermosensitive persons with multiple sclerosis. *Neurodegenerative Disease Management,* 6(1): 13-21.

Sandroff, B.M., (2015). Exercise and cognition in multiple sclerosis: The importance of acute exercise for developing better interventions. *Neuroscience & Biobehavioral Reviews,* 59: 173-183.

Schapiro, R.T. and Langer, S.L., (1994). Symptomatic therapy of multiple sclerosis. *Current Opinion in Neurology,* 7(3): 229-233.

Schulz, K.F., Altman, D.G., Moher, D., (2010). CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidimiology,* 63(8): 834-840.

Schwartz, M.H., Trost, J.P., Wervey, R.A., (2004). Measurement and management of errors in quantitative gait data. *Gait & Posture,* 20(2): 196-203.

Schwid, S.R., Goodman, A.D., McDermott, M.P., Bever, C.F., Cook, S.D., (2002). Quantitative functional measures in MS: what is a reliable change? *Neurology,* 58(8): 1294-1296.

Scott, S.M., van der Linden, M.L., Hooper, J.E., Cowan, P., Mercer, T.H., (2013). Quantification of gait kinematics and walking ability of people with multiple sclerosis who are new users of functional electrical stimulation. *Journal of Rehabilitation Medicine,* 45(4): 364-369.

Sechi, E., Keegan, B.M., Kaufmann, T.J., Kantarchi, O.H., Weinshenker, B.G., Flanagan, E.P., (2019). Unilateral motor progression in MS: Association with a critical corticospinal tract lesion. *Neurology*, 93(7): e628-e634.

Sehle, A., Vieten, M., Sailer, S., Mundermann, A., Dettmers, C., (2014). Objective assessment of motor fatigue in multiple sclerosis: the Fatigue index Kliniken Schmieder (FKS). *Journal of Neurology,* 261(9): 1752-1762.

Sehle, A., Mundermann, A., Starrost, K., Sailer, S., Becher, I., Dettmers, C., et al, (2011). Objective assessment of motor fatigue in multiple sclerosis using kinematic gait

analysis: a pilot study. *Journal of Neuroengineering and Rehabilitation,* 8:59. doi: 10.1186/1743-0003-8-59.

Sejbaek, T., Blaabjerg, M., Sprogoe, P., Ravnborg, M., (2018). Reliability and validity of a Danish version of the Multiple Sclerosis Neuropsychological Screening Questionnaire. *International Journal of MS Care,* 20(1): 49-54.

Severijns, D., Zijdewind, I., Dalgas, U., Lamers, I., Lismont, C., Feys, P., (2017). The assessment of motor fatigability in persons with multiple sclerosis: a systematic review. *Neurorehabilitation Neural Repair,* 31(5): 413-431.

Sharrack, B., Hughes, R.A., Soudain, S., Dunn, G., (1999). The psychometric properties of clinical rating scales used in multiple sclerosis. *Brain,* 122(Pt 1): 141-159.

Shrout, P.E. and Fleiss, J.L., (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin,* 86(2): 420-428.

Singh, S.J., Morgan, M.D., Scott, S., Walters, D., Hardman, A.E., (1992). Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax,* 47(12): 1019-1024.

Skjerbaek, A.G., Moller, A.B., Jensen, E., Vissing, K., Sorensen, H., Nybo, L., et al., (2013). Heat sensitive persons with multiple sclerosis are more tolerant to resistance exercise than to endurance exercise. *Multiple Sclerosis,* 19(7): 932-940.

Snook, E.M., Motl, R.W., Gliottoni, R.C., (2009). The effect of walking mobility on the measurement of physical activity using accelerometry in multiple sclerosis. *Clinical Rehabilitation,* 23(3): 248-258.

Socie, M.J., Motl, R.W., Sosnoff, J.J., (2014). Examination of spatiotemporal gait parameters during the 6-min walk in individuals with multiple sclerosis. *International Journal of Rehabilitation Research,* 37(4): 311-316.

Socie, M.J., Motl, R.W., Pula, J.H., Sandroff, B.M., Sosnoff, J.J., (2013). Gait variability and disability in multiple sclerosis. *Gait & Posture,* 38(1): 51-55.

Solaro,C., Brichetto, G., Amato, M.P., Cocco, E., Colombo, B., D'Aleo, G., et al., The prevalence of pain in multiple sclerosis: a multicentercross-sectional study. *Neurology,* 63(5):919-921.

Sonder, J.M., Mokkink, L.B., van der Linden, F.A., Polman, C.H., Uitdehaag, B.M., (2012). Validation and interpretation of the Dutch version of the Multiple Sclerosis Neuropsychological Screening Questionnaire. *Journal of the Neurological Sciences,* 320(1-2): 91-96.

Sorsdahl, A.B., Moe-Nilssen, R., Strand, L.I., (2008). Test-retest reliability of spatial and temporal gait parameters in children with cerebral palsy as measured by an electronic walkway. *Gait & Posture,* 27(1): 43-50.

Sosnoff, J.J., Klaren, R.E., Pilutti, L.A., Dlugonski, D., Motl, R.W., (2015). Reliability of gait in multiple sclerosis over 6 months. *Gait & Posture,* 41(3): 860-862.

Sosnoff, J.J., Socie, M.J., Sandroff, B.M., Balantrapu, S., Suh, Y., Pula, J.H., et al., (2014). Mobility and cognitive correlates of dual task cost of walking in persons with multiple sclerosis. *Disability Rehabilitation,* 36(3): 205-209.

Sosnoff, J.J., Sandroff, B.M., Motl, R.W., (2012). Quantifying gait abnormalities in persons with multiple sclerosis with minimal disability. *Gait & Posture,* 36(1): 154-156.

Sosnoff, J.J., Boes, M.K., Sandroff, B.M., Socie, M.J., Pula, J.H., Motl, R.W., (2011a). Walking and thinking in persons with multiple sclerosis who vary in disability. *Archives of Physical Medicine and Rehabilitation,* 92(12):2028-2033.

Sosnoff, J.J., Weikert, M., Dlugonski, D., Smith, D.C., Motl, R.W., (2011b). Quantifyig gait impairment in multiple sclerosis using GAITRite technology. *Gait & Posture,* 34(1): 145-147.

Steens A., de Vries, A., Hemmen, J., Heersema, T., Heerings, M., Maurits, N., et al., (2012). *Neurorehabilitation Neural Repair,* 26(1): 48-57.

Stein, R.B., Everaert, D.G., Thompson, A.J., Chong, S.L., Whittaker, M., Robertson, J., et al., (2010). Long-term therapeutic and orthotic effects of a foot drop stimulator on walking performance in progressive and nonprogressive neurological disorders. *Neurorehabilitation Neural Repair,* 24(2): 152-167.

Stein, R.B., Chong, S.L., Everaert, D.G., Rolf, R., Thompson, A.J., Whiattaker, M., et al., (2006). A multicenter trial of a footdrop stimulator controlled by a tilt sensor. *Neurorehabilitation Neural Repair,* 20(3): 371-379.

Steinwender, G., Saraph, V., Scheiber, S., Zwick, E.B., Uitz, C., Hackl, K., (2000). Intrasubject repeatability of gait analysis data in normal and spastic children. *Clinical Biomechanics (Bristol, Avon),* **15**(2): 134-139.

Stellmann, J.P., Neuhaus, A., Götze, N., Briken, S., Lederer, C., Schimpl, M., et al., (2015). Ecological validity of walking capacity tests in multiple sclerosis. *PLoS One,* 10(4): e0123822.

Stins, J.F. and Beek, P.J., (2012). A critical evaluation of the cognitive penetrability of posture. *Experimental Aging Research*, 38(2): 208-219.

Stokic, D.S., Horn, T.S., Ramshur, J.M., Chow, J.W., (2009). Agreement between temporospatial gait parameters of an electronic walkway and a motion capture system in healthy and chronic stroke populations. *American Journal of Physical Medicine & Rehabilitation,* 88(6): 437-444.

Street, T., Taylor, P., Swain, I., (2015). Effectiveness of functional electrical stimulation on walking speed, functional walking category, and clinically meaningful changes for people with multiple sclerosis. *Archives of Physical Medicine and Rehabilitation,* 96(4): 667-672.

Streiner, D.L. and Kottner, J., (2014). Recommendations for reporting the results of studies of instrument and scale development and testing. *Journal of Advanced Nursing,* 70(9): 1970-1979.

Sumowski, J.F. and Leavitt, V.M., (2014). Body temperature is elevated and linked to fatigue in relapsing-remitting multiple sclerosis, even without heat exposure. *Archives of Physical Medicine and Rehabilitation,* 95(7): 1298-1302.

Tang, K.T., Richardson, A.M., Maxwell, D., Spence, W.D., Stansfield, B.W., (2013). Evaluation of an activity monitor for the objective measurement of free-living physical activity in children with cerebral palsy. *Archives of Physical Medicine and Rehabilitation,* 94(12): 2549-2558.

Taylor, P.N., Humphreys, L., Swain, I., (2013). The long-term cost-effectiveness of the use of Functional Electrical Stimulation for the correction of dropped foot due to upper motor neuron lesion. *Journal of Rehabilitation Medicine,* 45(2): 154-160.

Taylor, P.N., Burridge, J.H., Dunkerley, A.L., Lamb A., Wood, D.E., Norton, J.A., et al., (1999). Patients' perceptions of the Odstock Dropped Foot Stimulator (ODFS). *Clinical Rehabilitation,* 13(5): 439-446.

Taylor, P.N., Burridge, J.H., Dunkerley, A.L., Wood, D.E., Norton, J.A., Singleton, C., et al., (1999). Clinical use of the Odstock dropped foot stimulator: its effect on the speed and effort of walking. *Archives of Physical Medicine and Rehabilitation,* 80(12): 1577-1583.

Terwee, C.B., Bot, S.D., de Boer, M.R., van der Windt, D.A., Knol, D.L., Dekker, J., et al., (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology,* 60(1): 34-42.

Terwee, C.B., Dekker, F.W., Wiersinga, W.M., Prummel, M.F., Bossuyt, P.M., (2003). On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Quality of Life Research,* 12(4): 349-362.

Terwee, C.B., Jansma, E.P., Riphagen, I.I., de Vet, H.C., (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research,* 18(8): 1115-1123.

Terwee, C.B., Mokkink, L.B., Knol, D.L., Ostelo, R.W., Bouter, L.M., de Vet, H.C., (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research,* 21(4): 651-657.

Terwee, C.B., Roorda, L.D., Knol, D.L., De Boer, M.R., de Vet, H.C., (2009). Linking measurement error to minimal important change of patient-reported outcomes. *Journal of Clinical Epidemiology,* 62(10): 1062-1067.

Thickbroom, G.W., Sacco, P., Faulkner, D.L., Kermode, A.G., Mastaglia, F.L., (2008). Enhanced corticomotor excitability with dynamic fatiguing exercise of the lower limb in multiple sclerosis. *Journal of Neurology,* 255(7): 1001-1005.

Toomey, E. and Coote, S., (2013). Between-rater reliability of the 6-minute walk test, berg balance scale, and handheld dynamometry in people with multiple sclerosis. *International Journal of MS Care,* 15(1): 1-6.

Trapp, B.D. and Nave, K.A., (2008). Multiple sclerosis: an immune or neurodegenerative disorder? *Annual Review of Neuroscience,* 31: 247-269.

van Bloemendaal, M., van de Water, A.T., van de Port, I.G., (2012). Walking tests for stroke survivors: a systematic review of their measurement properties. *Disability Rehabilitation,* 34(26): 2207-2221.

van der Linden, M.L., Andreopoulou, G., Scopes, J., Hooper, J.E., Mercer, T.H., (2018). Ankle Kinematics and Temporal Gait Characteristics over the Duration of a 6-Minute Walk Test in People with Multiple Sclerosis Who Experience Foot Drop. *Rehabilitation Research and Practice,* 2018: 1260852. doi: 10.1155/2018/1260852.

van der Linden, M.L., Hooper, J.E., Cowan, P., Weller, B.B., Mercer, T.H., (2014a). Habitual functional electrical stimulation therapy improves gait kinematics and walking performance, but not patient-reported functional outcomes, of people with multiple sclerosis who present with foot-drop. *PLoS One,* 9(8): e103368.

van der Linden, M.L., Scott, S.M., Hooper, J.E., Cowan, P., Mercer, T.H., (2014b). Gait kinematics of people with multiple sclerosis and the acute application of functional electrical stimulation. *Gait & Posture,* 39(4): 1092-1096.

van Tulder, M., Furlan, A., Bombardier, C., Bouter, L., (2003). Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine (Phila Pa 1976),* 28(12): 1290-1299.

van Winsen, L.M., Kragt, J.J., Hoogervorst, E.L., Polman, C.H., Uitdehaag, B.M., (2010). Outcome measurement in multiple sclerosis: detection of clinically relevant improvement. *Multiple Sclerosis,* 16(5): 604-610.

Verschuren, O., Takken, T., Ketelaar, M., Gorter, J.W., Helders, P.J., (2006). Reliability and validity of data for 2 newly developed shuttle run tests in children with cerebral palsy. *Physical Therapy,* 86(8): 1107-1117.

Wajda, D.A., Motl, R.W., Sosnoff, J.J., (2013). Dual task cost of walking is related to fall risk in persons with multiple sclerosis. *Journal of the Neurological Sciences,* 335(1-2): 160-163.

Wajda, D.A., Roeing, K.L., McAuley, E., Motl, R.W., Sosnoff, J.J., (2016). The Relationship Between Balance Confidence and Cognitive Motor Interference in Individuals With Multiple Sclerosis. *Journal of Motor Behavior,* 48(1): 66-71.

Wajda, D.A., Sandroff, B.M., Pula, J.H., Motl, R.W., Sosnoff, J.J., (2013). Effects of walking direction and cognitive challenges on gait in persons with multiple sclerosis. *Multiple Sclerosis International,* 2013: 859323. doi: 10.1155/2013/859323.

Wajda, D.A. and Sosnoff, J.J., (2015). Cognitive-motor interference in multiple sclerosis: a systematic review of evidence, correlates, and consequences. *Biomed Research International,* 2015: 720856. doi: 10.1155/2015/720856.

Wang, J.L., Reimer, M.A., Metz, L.M., Patten, S.B., (2000). Major depression and quality of life in individuals with multiple sclerosis. *International Journal of Psychiatry in Medicine,* 30(4): 309-317.

Watson, T.M., Ford, E., Worthington, E., Lincoln, N.B., (2014). Validation of mood measures for people with multiple sclerosis. *International Journal of MS Care,* 16(2): 105-109.

Weinshenker, B.G., Bass, B., Rice, G.P., Noseworthy, J., Carriere, W., Baskerville, J., et al., (1989). The natural history of multiple sclerosis: a geographically based study. 2. Predictive value of the early clinical course. *Brain,* 112 (Pt 6): 1419-1428.

Wening, J., Ford, J., Jouett, L.D., (2013). Orthotics and FES for maintenance of walking in patients with MS. *Disease-A-Month,* 59(8): 284-289.

Wilken, J.M., Rodriguez, K.M., Brawner, M., Darter, B.J., (2012). Reliability and Minimal Detectible Change values for gait kinematics and kinetics in healthy adults. *Gait & Posture,* 35(2): 301-307.

Witte, M.E., Mahad, D.J., Lassmann, H., van Horssen, J., (2014). Mitochondrial dysfunction contributes to neurodegeneration in multiple sclerosis. *Trends in Molecular Medicine,* 20(3): 179-187.

Wolkorte, R., Heersema, D.J., Zijdewind, I., (2015). Reduced Dual-Task Performance in MS Patients Is Further Decreased by Muscle Fatigue. *Neurorehabilitation Neural Repair*, 29(5): 424-435.

Yahia, A., Ghroubi, S., Mhiri, C., Elleuch, M.H., (2011). Relationship between muscular strength, gait and postural parameters in multiple sclerosis. *Annals of Physical and Rehabilitation Medicine,* 54(3): 144-155.

Yavuzer, G., Oken, O., Elhan, A., Stam, H.J., (2008). Repeatability of lower limb three-dimensional kinematics in patients with stroke. *Gait & Posture,* 27(1): 31-35.

Yogev-Seligmann, G., Rotem-Galili, Y., Disckstein, R., Giladi, N., Hausdorff, J.M., (2012). Effects of explicit prioritization on dual task walking in patients with Parkinson's disease. *Gait & Posture,* 35(4): 641-646.

Zigmond, A.S. and Snaith, R.P., (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica,* 67(6): 361-370.

Appendix 1: Foot drop audit

During annual clinical review of patients at the Anne Rowling Regenerative Medicine Clinic, we invited 80 ambulant pwMS to complete a short survey to explore their experience of exercise-related transient foot drop. Participants were included diagnosed with any subtype of MS and EDSS scores ranging from 0-5. Table 8 present the responses at each individual item of this questionnaire. Thirty-three of 47 respondents (70%) reported experiencing transient foot drop that comes during exercise, whilst three reported permanent foot drop and the remaining 11 did not report any foot drop. The most commonly reported activities associated with the onset of transient foot drop were walking (69%) and running (36%). The mean time for this to occur was 19 minutes of participation (range 2-60 minutes) or after approximately 1.65 miles. Based on the survey responses foot drop appears to be more evident, and occurs more quickly, during faster speeds of walking and running or walking uphill.

Table 8 Audit questionnaire presented as number of answers (%) or mean (range).

Questions	Answers
Q1. Do you experience temporarily 'foot drop' (after a certain period of walking or running or other type of activity? (n=47)	
Yes	33 (70%)
No, my foot drop starts immediately when I start walking/exercising	3 (6%)
No, I don't experience 'foot drop	11 (24%)

Q2. During which activity/activities do you experience this? (i.e. walking, running, cycling, driving) (n=36)	
Walking	25 (69%)
Running	13 (36%)
Cycling	3 (8%)
Other	4 (11%)
Q3. When do you start experiencing 'foot drop'? (i.e. after how many miles or minutes?) (n=36)	
Minutes	19 (2 - 60)
Miles	1.65 (0.5 - 4.5)
Q4. Does foot drop force you to stop your activity or reduce the intensity (i.e. walking/running speed) (n=34)	
I have to stop and rest	16 (47%)
I have to slow down/change my activity	18 (53%)
Q5. How long does it take approximately until you feel the foot drop has disappeared and you are able to resume walking/exercising or the activity you were doing? (n=36)	32 minutes (1-90)
Q6. Do you think your foot drop comes on earlier if: (n=36)	
You walk/run/cycle etc. faster	23 (64%)
You walk/run/cycle etc. uphill/upstairs	21(58%)
You walk/run on uneven terrain/crowded streets	18 (50%)
Your feel more fatigued than usual before your activity	22 (61%)
Q7. Do you think your foot drop comes on later if: (n=29)	
You walk/run/cycle etc. slower	18 (62%)
You walk/run etc. downhill/downstairs	10 (34%)
You walk/run etc. on a quiet, smooth level road	12 (41%)
Your feel less fatigued than usual before your activity	17 (58%)

Appendix 2: First search – overview of outcome measures

CINAHL search strategy/15.4.2016/ 1969-2016

S1. multiple sclerosis
S2. demyelinating disease
S3. demyelinating autoimmune disease
S4. chronic progressive multiple sclerosis
S5. secondary progressive multiple sclerosis
S6. primary progressive multiple sclerosis
S7. relapsing remitting multiple sclerosis
S8. clinically isolated syndrome
S9. demyelinating disorder
S10. transverse myelitis
S11. acute disseminated encephalomyelitis

S12. S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11

S13. functional electrical stimulation
S14. foot drop stimulation
S15. ankle foot orthosis
S16. splints
S17. electrical stimulation
S18. orthosis
S19. common peroneal stimulation
S20. electric stimulation
S21. peroneal nerve stimulation
S22. orthotic devices
S23. neuroprosthesis
S24. lower limb orthosis

S25. S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24

S26. S12 AND S25

MEDLINE search strategy/15.4.2016/1963-2016

S1. multiple sclerosis
S2. demyelinating disease
S3. demyelinating autoimmune disease
S4. chronic progressive multiple sclerosis
S5. secondary progressive multiple sclerosis

S6. primary progressive multiple sclerosis
S7. relapsing remitting multiple sclerosis
S8. clinically isolated syndrome
S9. demyelinating disorder
S10. transverse myelitis
S11. acute disseminated encephalomyelitis

S12. S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11

S13. functional electrical stimulation
S14. foot drop stimulation
S15. ankle foot orthosis
S16. splints
S17. electrical stimulation
S18. orthosis
S19. common peroneal stimulation
S20. electric stimulation
S21. peroneal nerve stimulation
S22. orthotic devices
S23. neuroprosthesis
S24. lower limb orthosis

S25. S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24

S26. S12 AND S25

SCOPUS search strategy /18.4.2016/1977-2016

(((TITLE-ABS-KEY (**multiple sclerosis**)) OR (TITLE-ABS-KEY (**demyelinating disease**)) OR (TITLE-ABS-KEY (**demyelinating autoimmune disease**)) OR (TITLE-ABS-KEY (**chronic progressive multiple sclerosis**)) OR (TITLE-ABS-KEY (**secondary progressive multiple sclerosis**)) OR (TITLE-ABS-KEY (**primary progressive multiple sclerosis**)) OR (TITLE-ABS-KEY (**relapsing remitting multiple sclerosis**)) OR (TITLE-ABS-KEY (**clinically isolated syndrome**))) OR ((TITLE-ABS-KEY (**demyelinating disorder**)) OR (TITLE-ABS-KEY (**transverse myelitis**)) OR (TITLE-ABS-KEY (**acute disseminated encephalomyelitis**)))) AND (((TITLE-ABS-KEY (**functional electrical stimulation**)) OR (TITLE-ABS-KEY (**foot drop stimulation**)) OR (TITLE-ABS-KEY (**ankle foot orthosis**)) OR (TITLE-ABS-KEY (**splints**)) OR (TITLE-ABS-KEY (**electrical stimulation**)) OR (TITLE-ABS-KEY (**orthosis**)) OR (TITLE-ABS-

KEY (**common peroneal stimulation**))) OR ((TITLE-ABS-
KEY (**electric stimulation**)) OR (TITLE-ABS-
KEY (**peroneal nerve stimulation**)) OR (TITLE-ABS-
KEY (**orthotic devices**)) OR (TITLE-ABS-
KEY (**neuroprosthesis**)) OR (TITLE-ABS-KEY (**lower limb orthotics**))))

Embase search strategy/15.4.2016/1974-2016

1. multiple sclerosis/
2. demyelinating disease/
3. demyelinating autoimmune disease.mp.
4. chronic progressive multiple sclerosis.mp.
5. secondary progressive multiple sclerosis.mp.
6. primary progressive multiple sclerosis.mp.
7. relapsing remitting multiple sclerosis.mp.
8. clinically isolated syndrome.mp.
9. demyelinating disorder.mp.
10. transverse myelitis.mp.
11. acute disseminated encephalomyelitis.mp. or acute disseminated encephalomyelitis/
12. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11
13. functional electrical stimulation.mp. or functional electrical stimulation/
14. foot drop stimulation.mp.
15. ankle foot orthosis.mp. or ankle foot orthosis/
16. splints.mp. or splint/
17. electrical stimulation.mp.
18. orthosis/ or orthosis.mp.
19. common peroneal stimulation.mp.
20. electric stimulation.mp.
21. peroneal nerve stimulation.mp.

22. orthotic devices.mp.

23. neuroprosthesis.mp. or neuroprosthesis/

24. lower limb orthosis.mp.

25. 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24

26. 12 and 25

Appendix 3: Principal search – systematic review of the psychometric properties of outcome measures

CINHAL/MEDLINE

1# Population
multiple sclerosis OR demyelinating disease OR demyelinating autoimmune disease OR chronic progressive multiple sclerosis OR secondary progressive multiple sclerosis OR primary progressive multiple sclerosis OR relapsing remitting multiple sclerosis OR clinically isolated syndrome OR demyelinating disorder OR transverse myelitis OR acute disseminated encephalomyelitis

2# Outcome measures
(3D gait analysis OR kinematics) OR (6MWT OR 6 min walk test OR six minute walking test) OR (walking speed OR walking speed over 10m) OR (Physiological Cost Index OR PCI) OR (effort of walking OR RPE OR Rate of Perceived Exertion) OR GAITRite OR standing balance test OR choice stepping reaction time OR (oxygen consumption OR oxygen uptake) OR (energy cost of walking OR mechanical energy) OR (MSFC OR Multiple Sclerosis Functional Composite) OR (T25FW OR Timed 25 Foot Walk) OR (mEFAP OR modified Emory Functional Ambulation Profile) OR (Rivermead Observational Gait Analysis OR ROGA) OROR falls diary OR (GAITRite FAP OR GAITRite Functional Ambulation Performance) OR Hauser Ambulation Index OR (2MWT OR 2 min walk test OR two minute walk test) OR (MSWS OR MS Walking Scale) OR (4MWT OR 4 min walk test OR four minute walk test) OR (3MWT OR 3 min walk test OR three minute walk test) OR EMG OR (stride parameters OR temporal parameters OR spatial parameters)

3# Filter search
instrumentation* OR methods OR validation stud* OR comparative stud* OR psychometr* OR clinimetr* OR clinometr* OR (MH "Outcomes (Health Care)+") OR (MH "Treatment Outcomes+") OR (MH “outcome assessment”) OR outcome assessment OR outcome measure* OR observer variation OR (MH “Health Status Indicators”) OR (MH “reproducibility of results”) OR reproducib* OR (MH “discriminant analysis”) OR reliab* OR unreliab* OR valid* OR coefficient OR homogeneity OR homogeneous OR “internal consistency”
OR (cronbach* AND (alpha OR alphas))
OR (item AND (correlation* OR selection* OR reduction*))
OR agreement OR precision OR imprecision OR “precise values” OR test-retest OR (test AND retest)
OR (reliab* AND (test OR retest))
OR stability OR interrater OR inter-rater OR intrarater OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver OR inter-observer OR intraobserver OR intra-observer OR intertechnician OR inter-technician OR intratechnician OR intra-technician OR interexaminer OR inter-examiner OR

intraexaminer OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant OR inter-participant OR intraparticipant OR intra-participant OR kappa OR kappa's OR kappas OR repeatab* OR ((replicab* OR repeated) AND (measure OR measures OR findings OR result OR results OR test OR tests))
OR generaliza* OR generalisa* OR concordance
OR (intraclass AND correlation*)
OR discriminative OR "known group" OR factor analysis OR factor analyses OR dimension* OR subscale*
OR (multitrait AND scaling AND (analysis OR analyses))
OR item discriminant OR interscale correlation* OR error OR errors OR "individual variability"
OR (variability AND (analysis OR values))
OR (uncertainty AND (measurement OR measuring))
OR "standard error of measurement" OR sensitiv* OR responsive*
OR ((minimal OR minimally OR clinical OR clinically) AND (important OR significant OR detectable) AND (change OR difference))
OR (small* AND (real OR detectable) AND (change OR difference))
OR meaningful change OR "ceiling effect" OR "floor effect" OR "Item response model" OR IRT OR Rasch OR "Differential item functioning" OR DIF OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")

5# Combination
1# AND 2# AND 3#

SCOPUS

1# POPULATION
(TITLE-ABS-KEY(multiple sclerosis)) OR (TITLE-ABS-KEY(demyelinating disease)) OR(TITLE-ABS-KEY(demyelinating autoimmune disease)) OR (TITLE-ABS-KEY(chronic progressive multiple sclerosis)) OR (TITLE-ABS-KEY(secondary progressive multiple sclerosis)) OR (TITLE-ABS-KEY(primary progressive multiple sclerosis)) OR (TITLE-ABS-KEY(relapsing remitting multiple sclerosis)) OR (TITLE-ABS-KEY(clinically isolated syndrome)) OR (TITLE-ABS-KEY(demyelinating disorder)) OR (TITLE-ABS-KEY(transverse myelitis)) OR (TITLE-ABS-KEY(acute disseminated encephalomyelitis))

2#OUTCOME MEASURES
(TITLE-ABS-KEY(3D gait analysis OR kinematics)) OR (TITLE-ABS-KEY(6MWT OR 6 min walk test OR six minute walk test)) OR (TITLE-ABS-KEY(walking speed OR walking speed over 10m)) OR (TITLE-ABS-KEY(Physiological Cost Index OR PCI)) OR (TITLE-ABS-KEY(effort of walking OR RPE OR Rate of Perceived

Exertion)) OR (TITLE-ABS-KEY(GAITRite)) OR (TITLE-ABS-KEY(standing balance test)) OR (TITLE-ABS-KEY(choice stepping reaction time)) OR (TITLE-ABS-KEY(oxygen consumption OR oxygen uptake)) OR (TITLE-ABS-KEY(energy cost of walking OR mechanical energy)) OR (TITLE-ABS-KEY(MSFC OR Multiple Sclerosis Functional Composite)) OR (TITLE-ABS-KEY(T25FW OR Timed 25 Foot Walk)) OR (TITLE-ABS-KEY(mEFAP OR modified Emory Functional Ambulation Profile)) OR (TITLE-ABS-KEY(Rivermead Observational Gait Analysis OR ROGA)) OR (TITLE-ABS-KEY(GAITRite FAP OR GAITRite Functional Ambulation Performance)) OR (TITLE-ABS-KEY(Hauser Ambulation Index)) OR (TITLE-ABS-KEY(2MWT OR 2 min walk test OR two minute walk test)) OR (TITLE-ABS-KEY(MSWS OR MS Walking Scale)) OR (TITLE-ABS-KEY(4MWT OR 4 min walk test OR four minute walk test)) OR (TITLE-ABS-KEY(3MWT OR 3 min walk test OR three minute walk test)) OR (TITLE-ABS-KEY(EMG)) OR (TITLE-ABS-KEY(stride parameters OR temporal parameters OR spatial parameters))

3# FILTER/SENSITIVE

instrumentation* OR methods OR validation stud* OR comparative stud* OR psychometr* OR clinimetr* OR clinometr* OR ("Outcomes (Health Care)+") OR ("Treatment Outcomes+") OR ("outcome assessment") OR outcome assessment OR outcome measure* OR observer variation OR ("Health Status Indicators") OR ("reproducibility of results") OR reproducib* OR ("discriminant analysis") OR reliab* OR unreliab* OR valid* OR coefficient OR homogeneity OR homogeneous OR "internal consistency" OR (cronbach* AND (alpha OR alphas)) OR (item AND (correlation* OR selection* OR reduction*)) OR agreement OR precision OR imprecision OR "precise values" OR test-retest OR (test AND retest) OR (reliab* AND (test OR retest)) OR stability OR interrater OR inter-rater OR intrarater OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver OR inter-observer OR intraobserver OR intra-observer OR intertechnician OR inter-technician OR intratechnician OR intra-technician OR interexaminer OR inter-examiner OR intraexaminer OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant OR inter-participant OR intraparticipant OR intra-participant OR kappa OR kappa's OR kappas OR repeatab* OR ((replicab* OR repeated) AND (measure OR measures OR findings OR result OR results OR test OR tests)) OR generaliza* OR generalisa* OR concordance OR (intraclass AND correlation*) OR discriminative OR "known group" OR factor analysis OR factor analyses OR dimension* OR subscale* OR (multitrait AND scaling AND (analysis OR analyses)) OR item discriminant OR interscale correlation* OR error OR errors OR "individual variability" OR (variability AND (analysis OR values)) OR (uncertainty AND (measurement OR measuring)) OR "standard error of measurement" OR sensitiv* OR responsive* OR ((minimal OR minimally OR clinical OR clinically) AND (important OR significant OR detectable) AND (change OR difference)) OR (small* AND (real OR detectable) AND (change OR difference)) OR meaningful change OR "ceiling effect" OR "floor effect" OR "Item Response Model" OR IRT OR Rasch OR "Differential item functioning" OR DIF OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence"

EMBASE

1# POPULATION

1. multiple sclerosis/
2. demyelinating disease/
3. demyelinating autoimmune disease.mp.
4. chronic progressive multiple sclerosis.mp.
5. secondary progressive multiple sclerosis.mp.
6. primary progressive multiple sclerosis.mp.
7. relapsing remitting multiple sclerosis.mp.
8. clinically isolated syndrome.mp.
9. demyelinating disorder.mp.
10. transverse myelitis.mp.
11. acute disseminated encephalomyelitis.mp. or acute disseminated encephalomyelitis/
12. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11

2# OUTCOME MEASURES

(3D gait analysis or kinematics).mp.
OR

(6MWT or 6 min walk test or six minute walk test).mp
OR (walking speed or walking speed over 10m).mp.

OR

(Physiological Cost Index or PCI).mp.

OR (effort of walking or RPE or Rate of Perceived Exertion).mp. OR GAITRite.mp.

OR standing balance test.mp. OR choice stepping reaction time.mp. OR (oxygen consumption or oxygen uptake).mp. OR (energy cost of walking or mechanical energy).mp. OR (MSFC or Multiple Sclerosis Functional Composite).mp. OR (T25FW

or Timed 25 Foot Walk).mp. OR (mEFAP or modified Emory Functional Ambulation Profile).mp. OR (Rivermead Observational Gait Analysis or ROGA).mp. OR (GAITRite FAP or GAITRite Functional Ambulation Performance).mp. OR Hauser Ambulation Index.mp. OR (2MWT or 2 min walk test or two minute walk test).mp. OR (MSWS or MS Walking Scale).mp. OR (4MWT or 4 min walk test or four minute walk test).mp. OR (3MWT or 3 min walk test or three minute walk test).mp. OR EMG.mp. OR (stride parameters or temporal parameters or spatial parameters).mp.

3# FILTER/SENSITIVE

('validation study'/ OR psychometr*.mp. OR clinimetr*.mp. OR 'outcome assessment'/ OR outcome assessment*.mp. OR 'comparative study'/ OR outcome measure*.mp. OR 'observer variation'/ OR 'observer variation'.mp OR 'reproducibility'/ OR reproducib*.mp. OR 'discriminant analysis'/ OR reliab*.mp. OR unreliab*.mp. OR valid*.mp. OR coefficient.mp. OR homogeneity.mp. OR homogeneous.mp. OR 'internal consistency'.mp. OR (cronbach* AND (alpha OR alphas)) OR (item AND (correlation* OR selection* OR reduction*)) OR agreement.mp. OR precision.mp. OR imprecision.mp. OR 'precise values'.mp. OR test-retest.mp. OR (test AND retest).mp. OR (reliab* AND (test OR retest)).mp. OR stability.mp. OR interrater.mp. OR inter-rater.mp. OR intrarater.mp. OR intra-rater.mp. OR intertester.mp. OR inter-tester.mp. OR intratester.mp. OR intra-tester.mp. OR interobserver.mp. OR inter-observer.mp. OR intraobserver.mp. OR intra-observer.mp. OR intertechnician.mp. OR inter-technician.mp. OR intratechnician.mp. OR intra-technician.mp. OR interexaminer.mp. OR inter-examiner.mp. OR intraexaminer.mp. OR intra-examiner.mp. OR interassay.mp. OR inter-assay.mp. OR intraassay.mp. OR intra-assay.mp. OR interindividual.mp. OR inter-individual.mp. OR intraindividual.mp. OR intra-individual.mp. OR interparticipant.mp. OR inter-participant.mp. OR intraparticipant.mp. OR intra-participant.mp. OR kappa.mp. OR kappa*.mp. OR kappas.mp. OR repeatab*.mp. OR ((replicab* OR repeated) AND (measure OR measures OR findings OR result OR results OR test OR tests)).mp. OR generaliza*.mp. OR generalisa*.mp. OR concordance.mp. OR (intraclass AND correlation*).mp. OR discriminative.mp. OR 'known group'.mp. OR factor analysis.mp. OR factor analyses.mp. OR dimension*.mp. OR subscale*.mp. OR (multitrait AND scaling AND (analysis OR analyses)).mp. OR item discriminant.mp. OR interscale correlation*.mp. OR (error/ OR error.mp.) OR errors.mp. OR 'individual variability'.mp. OR (variability AND (analysis OR values)).mp. OR (uncertainty AND (measurement OR measuring)).mp. OR 'standard error of measurement'.mp. OR sensitiv*.mp OR responsive* OR ((minimal OR minimally OR clinical OR clinically) AND (important OR significant OR detectable) AND (change OR difference)).mp. OR (small* AND (real OR detectable) AND (change OR difference)).mp. OR meaningful change.mp. OR 'ceiling effect'.mp. OR 'floor

effect’.mp. OR ‘Item response model’.mp. OR IRT.mp. OR Rasch.mp. OR ‘Differential item functioning’.mp. OR DIF.mp. OR ‘computer adaptive testing’.mp. OR ’item bank’.mp. OR ‘cross-cultural equivalence’.mp.)

Appendix 4: COSMIN checklist

COSMIN checklist with 4-point scale

Contact
CB Terwee, PhD
VU University Medical Center
Department of Epidemiology and Biostatistics
EMGO Institute for Health and Care Research
1081 BT Amsterdam
The Netherlands
Website: www.cosmin.nl, www.emgo.nl
E-mail: cb.terwee@vumc.nl

Instructions

This version of the COSMIN checklist is recommended for use in systematic reviews of measurement properties. With this version it is possible to calculate overall methodological quality scores per study on a measurement property. A methodological quality score per box is obtained by taking the lowest rating of any item in a box ('worse score counts'). For example, if for a reliability study one item in the box 'Reliability' is scored poor, the methodological quality of that reliability study is rated as poor. The Interpretability box and the Generalizability box are mainly used as data extraction forms. We recommend to use the Interpretability box to extract all information on the interpretability issues described in this box (e.g. norm scores, floor-ceiling effects, minimal important change) of the instruments under study from the included articles. Similar, we recommend to use the Generalizability box to extract data on the characteristics of the study population and sampling procedure. Therefore no scoring system was developed for these boxes.

This scoring system is described in this paper:

Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on

measurement properties: a scoring system for the COSMIN checklist. Quality of Life Research 2011, July 6 [epub ahead of print].

Step 1. Evaluated measurement properties in the article

	Internal consistency	Box A
	Reliability	Box B
	Measurement error	Box C
	Content validity	Box D
	Structural validity	Box E
	Hypotheses testing	Box F
	Cross-cultural validity	Box G
	Criterion validity	Box H
	Responsiveness	Box I

Step 2. Determining if the statistical method used in the article are based on CTT or IRT

Box General requirements for studies that applied Item Response Theory (IRT) models				
	excellent	**good**	**fair**	**poor**
1 Was the IRT model used adequately described? e.g. One Parameter Logistic Model (OPLM), Partial Credit Model (PCM), Graded Response Model (GRM)	IRT model adequately described	IRT model not adequately described		
2 Was the computer software package used adequately described? e.g. RUMM2020, WINSTEPS, OPLM, MULTILOG, PARSCALE, BILOG, NLMIXED	Software package adequately described	Software package not adequately described		
3 Was the method of estimation used adequately described? e.g. conditional maximum likelihood (CML), marginal maximum likelihood (MML)	Method of estimation adequately described	Method of estimation not adequately described		
4 Were the assumptions for estimating parameters of the IRT model checked? e.g. unidimensionality, local independence, and item fit (e.g. differential item functioning (DIF))	assumptions of the IRT model checked	assumptions of the IRT model partly checked	assumptions of the IRT model not checked or unknown	

To obtain a total score for the methodological quality of studies that use IRT methods, the 'worse score counts' algorithm should be applied to the IRT box in combination with the box of the measurement property that was evaluated in the IRT study. For example, if IRT methods are used to study internal consistency and item 4 in the IRT box is scored fair, while the items in the internal consistency box (box A) are all scored as

good or excellent, the methodological quality score for internal consistency will be fair. However, if any of the items in box A is scored poor, the methodological quality score for internal consistency will be poor.

Step 3. Determining if a study meets the standards for good methodological quality

Box A. Internal consistency				
	excellent	**Good**	**fair**	**poor**
1 Does the scale consist of effect indicators, i.e. is it based on a reflective model?				
Design requirements				
2 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
3 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
4 Was the sample size included in the internal consistency analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
5 Was the unidimensionality of the scale checked? i.e. was factor analysis or IRT model applied?	Factor analysis performed in the study population	Authors refer to another study in which factor analysis was performed in a similar study population	Authors refer to another study in which factor analysis was performed, but not in a similar study population	Factor analysis NOT performed and no reference to another study

6	Was the sample size included in the unidimensionality analysis adequate?	7* #items and ≥100	5* #items and ≥100 OR 6-7* #items but <100	5* #items but <100	<5* #items

7	Was an internal consistency statistic calculated for each (unidimensional) (sub)scale separately?	Internal consistency statistic calculated for each subscale separately			Internal consistency statistic NOT calculated for each subscale separately
8	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods					
9	for Classical Test Theory (CTT), continuous scores: Was Cronbach's alpha calculated?	Cronbach's alpha calculated		Only item-total correlations calculated	No Cronbach's alpha and no item-total correlations calculated
10	for CTT, dichotomous scores: Was Cronbach's alpha or KR-20 calculated?	Cronbach's alpha or KR-20 calculated		Only item-total correlations calculated	No Cronbach's alpha or KR-20 and no item- total correlations calculated
11	for IRT: Was a goodness of fit statistic at a global level calculated? E.g. χ^2, reliability coefficient of estimated latent trait value (index of (subject or item) separation)	Goodness of fit statistic at a global level calculated			Goodness of fit statistic at a global level NOT calculated

NB. Item 1 is used to determine whether internal consistency is relevant for the instrument under study. It is not used to rate the quality of the study.

Box B. Reliability: relative measures (including test-retest reliability, inter-rater reliability and intra-rater reliability)				
	excellent	**Good**	**fair**	**poor**
Design requirements				
1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3 Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4 Were at least two measurements available?	At least two measurements			Only one measurement
5 Were the administrations independent?	Independent measurements	Assumable that the measurements were independent	Doubtful whether the measurements were independent	measurements NOT independent
6 Was the time interval stated?	Time interval stated		Time interval NOT stated	
7 Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable

8	Was the time interval appropriate?	Time interval appropriate	Doubtful whether time interval was appropriate Time interval NOT appropriate

9	Were the test conditions similar for both measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar
10	Were there any important flaws in the design or methods of the study? methodological flaws in the design or execution of the study	No other important		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods					
11	for continuous scores: Was an intraclass correlation coefficient (ICC) calculated? and model or formula of the ICC is described	ICC calculated	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated
12	for dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			Only percentage agreement calculated
13	for ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated	Only percentage agreement

calculated

14	for ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described

Box C. Measurement error: absolute measures				
	excellent	**Good**	**fair**	**poor**
Design requirements				
1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3 Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4 Were at least two measurements available?	At least two measurements			Only one measurement
5 Were the administrations independent?	Independent measurements	Assumable that the measurements were independent	Doubtful whether the measurements were independent	measurements NOT independent
6 Was the time interval stated?	Time interval stated		Time interval NOT stated	
7 Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable

8	Was the time interval appropriate?	Time interval appropriate	Doubtful whether time interval was appropriate	Time interval NOT appropriate

9	Were the test conditions similar for both measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar
10	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods					
11	for CTT: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population

Box D. Content validity (including face validity)				
	excellent	**Good**	**fair**	**poor**
General requirements				
1 Was there an assessment of whether all items refer to relevant aspects of the construct to be measured?	Assessed if all items refer to relevant aspects of the construct to be measured		Aspects of the construct to be measured poorly described AND this was not taken into consideration	NOT assessed if all items refer to relevant aspects of the construct to be measured

2	Was there an assessment of whether all items are relevant for the study population? (e.g. age, gender, disease characteristics, country, setting)	Assessed if all items are relevant for the study population in adequate sample size (≥10)	Assessed if all items are relevant for the study population in moderate sample size (5-9)	Assessed if all items are relevant for the study population in small sample size (<5)	NOT assessed if all items are relevant for the study population OR target population not involved
3	Was there an assessment of whether all items are relevant for the purpose of the measurement instrument? (discriminative, evaluative, and/or predictive)	Assessed if all items are relevant for the purpose of the application	Purpose of the instrument was not described but Assumed	NOT assessed if all items are relevant for the purpose of the application	
4	Was there an assessment of whether all items together comprehensively reflect the construct to be measured?	Assessed if all items together comprehensively reflect the construct to be measured		No theoretical foundation of the construct and this was not taken into consideration	NOT assessed if all items together comprehen-sively reflect the construct to be measured
5	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study

Box E. Structural validity				
	excellent	**Good**	**fair**	**poor**
1 Does the scale consist of effect indicators, i.e. is it based on a reflective model?				
Design requirements				
2 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
3 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
4 Was the sample size included in the analysis adequate?	7* #items and ≥100	5* #items and ≥100 OR 5-7* #items but <100	5* #items but <100	<5* #items
5 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study (e.g. rotation method not described)	Other important methodological flaws in the design or execution of the study (e.g. inappropriate rotation method)

Statistical methods			
6 for CTT: Was exploratory or confirmatory factor analysis performed?	Exploratory or confirmatory factor analysis performed and type of factor analysis appropriate in view of existing information	Exploratory factor analysis performed while confirmatory would have been more appropriate	No exploratory or confirmatory factor analysis performed
7 for IRT: Were IRT tests for determining the (uni-) dimensionality of the items performed?	IRT test for determining (uni)dimension- ality performed		IRT test for determining (uni)dimension- ality NOT performed

Box F. Hypotheses testing				
	excellent	**Good**	**fair**	**Poor**
Design requirements				
1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	

3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100 per analysis)	Good sample size (50-99 per analysis)	Moderate sample size (30-49 per analysis)	Small sample size (<30 per analysis)

4	Were hypotheses regarding correlations or mean differences formulated a priori (i.e. before data collection)?	Multiple hypotheses formulated a priori	Minimal number of hypotheses formulate a priori	Hypotheses vague or not formulated but possible to deduce what was expected	Unclear what was expected
5	Was the expected *direction* of correlations or mean differences included in the hypotheses?	Expected direction of the correlations or differences stated	Expected direction of the correlations or differences NOT stated		
6	Was the expected absolute or relative *magnitude* of correlations or mean differences included in the hypotheses?	Expected magnitude of the correlations or differences stated	Expected magnitude of the correlations or differences NOT stated		
7	for convergent validity: Was an adequate description provided of the comparator instrument(s)?	Adequate description of the constructs measured by the comparator instrument(s)	Adequate description of most of the constructs measured by the comparator instrument(s)	Poor description of the constructs measured by the comparator instrument(s)	NO description of the constructs measured by the comparator instrument(s)
8	for convergent validity: Were the measurement properties of the comparator instrument(s) adequately described?	Adequate measurement properties of the comparator instrument(s) in a population similar to the study population	Adequate Measurement properties of the Comparator instrument(s) but not sure if these apply to the study Population	Some information on measurement properties (or a reference to a study on measurement properties) of the comparator instrument(s) in any study	No information on the measurement properties of the comparator instrument(s)

	population

9 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study (e.g. only data presented on a comparison with an instrument that measures another construct)	Other important methodological flaws in the design or execution of the study
Statistical methods				
10 Were design and statistical methods adequate for the hypotheses to be tested?	Statistical methods applied appropriate	Assumable that statistical methods were appropriate, e.g. Pearson Correlations applied, but distribution of scores or mean (SD) not Presented	Statistical methods applied NOT optimal	Statistical methods applied NOT appropriate

Box G. Cross-cultural validity				
	excellent	**Good**	**fair**	**poor**
Design requirements				
1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		

2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled

3	Was the sample size included in the analysis adequate?	CTT: 7* #items and ≥100 IRT: ≥200 per group	CTT: 5* #items and ≥100 OR 5-7* #items but <100 IRT: ≥200 in 1 group and 100-199 in 1 group	CTT: 5* #items but <100 IRT: 100-199 per group	CTT: <5* #items IRT: (<100 in 1 or both groups
4	Were both the original language in which the HR-PRO instrument was developed, and the language in which the HR-PRO instrument was translated described?	Both source language and target language described			Source language NOT known
5	Was the expertise of the people involved in the translation process adequately described? e.g. expertise in the disease(s) involved, expertise in the construct to be measured, expertise in both languages	Expertise of the translators described with respect to disease, construct, and language	Expertise of the translators with respect to disease or construct poor or not described	Expertise of the translators with respect to language not described	
6	Did the translators work independently from each other?	Translators worked independent	Assumable that the translators worked independent	Unclear whether translators worked independent	Translators worked NOT independent
7	Were items translated forward and backward?	Multiple forward and multiple backward translations	Multiple forward translations but one backward translation	One forward and one backward translation	Only a forward translation

8	Was there an adequate description of how differences between the original and translated versions were resolved?	Adequate description of how differences between translators were resolved	Poorly or NOT described how differences between translators were resolved

9	Was the translation reviewed by a committee (e.g. original developers)?	Translation reviewed by a committee (involving other people than the translators, e.g. the original developers)	Translation NOT reviewed by (such) a Committee		
10	Was the HR-PRO instrument pre-tested (e.g. cognitive interviews) to check interpretation, cultural relevance of the translation, and ease of comprehension?	Translated instrument pre-tested in the target population	Translated instrument pre-tested, but unclear if this was done in the target Population	Translated instrument pre-tested, but NOT in the target population	Translated instrument NOT pre-tested
11	Was the sample used in the pre-test adequately described?	Sample used in the pre-test adequately described		Sample used in the pre-test NOT (adequately) described	
12	Were the samples similar for all characteristics except language and/or cultural background?	Shown that samples were similar for all characteristics except language /culture	Stated (but not shown) that samples were similar for all Characteristics except language /culture	Unclear whether samples were similar for all characteristics except language /culture	Samples were NOT similar for all characteristics except language /culture
13	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study

Statistical methods		
14 for CTT: Was confirmatory factor analysis performed?	Multiple-group confirmatory factor analysis performed	Multiple-group confirmatory factor analysis NOT performed
15 for IRT: Was differential item function (DIF) between language groups assessed?	DIF between language groups assessed	DIF between language groups NOT assessed

Box H. Criterion validity				
	excellent	**Good**	**fair**	**poor**
Design requirements				
1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3 Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)

4	Can the criterion used or employed be considered as a reasonable ‘gold standard’?	Criterion used can be considered an adequate ‘gold standard’ (evidence provided)	No evidence provided, but assumable that the criterion used can be considered an adequate ‘gold standard’	Unclear whether the criterion used can be considered an adequate ‘gold standard’	Criterion used can NOT be considered an adequate ‘gold standard’

5 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study	Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods			
6 for continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated		Correlations or AUC NOT calculated
7 for dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated		Sensitivity and specificity NOT calculated

Box I. Responsiveness				
	excellent	**Good**	**fair**	**poor**
Design requirements				
1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3 Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4 Was a longitudinal design with at least two measurement used?	Longitudinal design used			No longitudinal design used

5 Was the time interval stated?	Time interval adequately described	Time interval NOT described

6	If anything occurred in the interim period (e.g. intervention, other relevant events), was it adequately described?	Anything that occurred during the interim period (e.g. treatment) adequately described	Assumable what occurred during the interim period	Unclear or NOT described what occurred during the interim period	
7	Was a proportion of the patients changed (i.e. improvement or deterioration)?	Part of the patients were changed (evidence provided)	NO evidence provided, but assumable that part of the patients were changed	Unclear if part of the patients were changed	Patients were NOT changed
Design requirements for hypotheses testing					
For constructs for which a gold standard was not available:					
8	Were hypotheses about changes in scores formulated a priori (i.e. before data collection)?	Hypotheses formulated a priori		Hypotheses vague or not formulated but possible to deduce what was expected	Unclear what was expected
9	Was the expected *direction* of correlations or mean differences of the change scores of HR-PRO instruments included in these hypotheses?	Expected direction of the correlations or differences stated	Expected direction of the correlations or differences NOT stated		
10	Were the expected absolute or relative *magnitude* of correlations or mean differences of the change scores of HR-PRO instruments included in these hypotheses?	Expected magnitude of the correlations or differences stated	Expected magnitude of the correlations or differences NOT stated		

11 Was an adequate description provided of the comparator instrument(s)?	Adequate description of the constructs measured by the comparator instrument(s)		Poor description of the constructs measured by the comparator instrument(s)	NO description of the constructs measured by the comparator instrument(s)
12 Were the measurement properties of the comparator instrument(s) adequately described?	Adequate measurement properties of the comparator instrument(s) in a population similar to the study population	Adequate measurement properties of the comparator instrument(s) but not sure if these apply to the study population	Some information on measurement properties (or a reference to a study on measurement properties) of the comparator instrument(s) in any study population	NO information on the measurement properties of the comparator instrument(s)
13 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study (e.g. only data presented on a comparison with an instrument that measures another construct)	Other important methodological flaws in the design or execution of the study
Statistical methods				
14 Were design and statistical methods adequate for the hypotheses to be tested?	Statistical methods applied appropriate		Statistical methods applied NOT optimal	Statistical methods applied NOT appropriate

Design requirement for comparison to a gold standard				
For constructs for which a gold standard was available:				
15 Can the criterion for change be considered as a reasonable gold standard?	Criterion used can be considered an adequate 'gold standard' (evidence provided)	No evidence provided, but assumable that the criterion used can be considered an adequate 'gold standard'	Unclear whether the criterion used can be considered an adequate 'gold standard'	Criterion used can NOT be considered an adequate 'gold standard'
16 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods				
17 for continuous scores: Were correlations between change scores, or the area under the Receiver Operator Curve (ROC) curve calculated?	Correlations or Area under the ROC Curve (AUC) calculated			Correlations or AUC NOT calculated
18 for dichotomous scales: Were sensitivity and specificity (changed versus not changed) determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated

Interpretability

We recommend to use the Interpretability box to extract all information on the interpretability issues described in this box of the instruments under study from the included articles.

Box Interpretability	
Percentage of missing items	
Description of how missing items were handled	
Distribution of the (total) scores	
Percentage of the respondents who had the lowest possible (total) score	
Percentage of the respondents who had the highest possible (total) score	
Scores and change scores (i.e. means and SD) for relevant (sub) groups, e.g. for normative groups, subgroups of patients, or the general population	
Minimal Important Change (MIC) or Minimal Important Difference (MID)	

Generalizability

We recommend to use the Generalizability box to extract data on the characteristics of the study populations and sampling procedures of the included studies.

Box Generalisability	
Median or mean age (with standard deviation or range)	
Distribution of sex	
Important disease characteristics (e.g. severity, status, duration) and description of treatment	
Setting(s) in which the study was conducted (e.g. general population, primary care or hospital/rehabilitation care)	
Countries in which the study was conducted	
Language in which the HR-PRO instrument was evaluated	
Method used to select patients (e.g. convenience, consecutive, or random)	
Percentage of missing responses (response rate)	

Appendix 5: Guidelines for rating COSMIN

1. If there are only references for the psychometric properties of the comparator instruments, we rate them as fair.
2. If it's a within-day testing, we rate as excellent the questions of being the patients stable & the interval time between testing.
3. In Box I-12 if the paper talks about global scales, we rate it as n/a.
4. In Box I-8 if we rate it as poor, then the 9 & 10 are n/a because there is not a hypothesis formulated so that there exists a direction or a magnitude.
5. In Box-B if they describe ICC for continues scores, we rate as n/a the 12-14 questions.
6. If hypothesis, magnitudes and directions are formulated in methods section are still rated as excellent.
7. If the instrument is not a questionnaire, then rate as n/a the questions of missing items given & how they were handled.
8. EDSS take it always as a 'gold standard' – BUT MENTION IN DISCUSSION THAT ITS NOT
9. EDSS take it as a gold standard only for comparison with walking ability scales and rate it as 'fair'.
10. In Box – I, every paper that don't mentions what happened in the interim period or mention if the patients changed we rate it as fair.
11. When we rate the interim period and the OM is a questionnaire, it has to be at least 7 days between each measurement, so that the time will be appropriate (memory of the previous answers?).

Guidelines between reviewers for rating COSMIN	
When there are only references for the psychometric properties of the comparator instrument, we rate them as 'fair'	
EDSS we take it always as a 'gold standard', but we rate it as 'fair'	
Box B/C	QSN 7 & 8: if it's a within-day testing, we rate them as 'excellent' & if the instrument is a questionnaire it has to be at least 7 days to rate them as 'excellent'
Box B/C/F/H/I	QSN 1 & 2: if the instrument is not a questionnaire, we rate these questions as 'n/a'
Box B	QSN 11: if the ICC is calculated for continuous scores, we rate QSN 12-14 as 'n/a'
Box F Box I	QSN 4-6: if they are mentioned in methods, we rate them as 'excellent' QSN 8-10: if they are mentioned in methods, we rate them as 'excellent'
Box I	QSN 6-7: if they don't mention what happened & if patients were stable, we rate them as 'fair'
	QSN 8: if this question is 'poor', then questions 9-10 are 'n/a' because magnitude & direction cannot exist if there is no hypothesis
	QSN 12: if the comparator instrument is a global rating scale, we rate it as 'n/a'

Appendix 6: Ethical opinion from the National Health Service (NHS) Research Ethics Committee (Chapter 4 & 5)

Lothian NHS Board

South East Scotland Research Ethics Committee 02

Waverley Gate
2-4 Waterloo Place
Edinburgh
EH1 3EG
Telephone 0131 536 9000

www.nhslothian.scot.nhs.uk

Direct Line: 0131 465 5674
Email: Joyce.Clearie@nhslothian.scot.nhs.uk
Enquiries to: Joyce Clearie
Extension: 35674

Please note: This is the favourable opinion of the REC only and does not allow the amendment to be implemented at NHS sites in England until the outcome of the HRA assessment has been confirmed.

01 November 2016

Dr Don Mahad
University of Edinburgh
Centre for Neuroregeneration
Chancellor's Building
49 Little France Crescent
Edinburgh, EH16 4SB

Dear Dr Mahad,

Study title:	**Performance related fatigue and disease progression in multiple sclerosis**
REC reference:	**15/SS/0088**
Protocol number:	**N/A**
Amendment number:	**REC Ref 15/SS/0088/AM01**
Amendment date:	**15 October 2016**
IRAS project ID:	**166909**

The above amendment was reviewed on 26 October 2016 by the Sub-Committee in correspondence.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

The Committee had no ethical concerns regarding this amendment.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Non-validated questionnaire [FSMC Questionnaire]	1.0	
Notice of Substantial Amendment (non-CTIMP)		15 October 2016
Participant information sheet (PIS) [Healthy volunteer with highlighted changes]	4.0	15 October 2016

Appendix 7: NHS Lothian Research and Development approval (Chapter 4 & 5)

University Hospitals Division

NHS Lothian

Queen's Medical Research Institute
47 Little France Crescent, Edinburgh, EH16 4TJ

FM/LM/approval

12 August 2016

Dr Don Mahad
University of Edinburgh
49 Little France Crescent
Chancellor's Building
Edinburgh
EH16 4SB

Research & Development
Room E1.12
Tel: 0131 242 3330

Email:
accord@nhslothian.scot.nhs.uk

Director: Professor David E Newby

Dear Dr Mahad

Lothian R&D Project No: 2016/0234 **REC No:** 15/SS/0088

Title of Research: Performance related fatigue and disease progression in multiple sclerosis

Participant Information Sheet (Patients, Healthy Volunteers):
Version 3.0 dated 15 April 2015

Consent Form:
Version 3.0 dated 15 April 2015

Protocol: Version 3.0 dated 15 April 2015

I am pleased to inform you this letter provides Site Specific approval for NHS Lothian for the above study and you may proceed with your research, subject to the conditions below.

We note that this project includes a researcher(s) who will require a Letter of Access from NHS Lothian. The individual(s) concerned (Georgia Andreopoulou) should contact our offices with a view to applying for the necessary documentation. Please note all final paperwork will have to be signed and returned to our R&D offices before the researcher(s) can commence work on the project.

Please note that the NHS Lothian R&D Office must be informed of any changes to the study such as amendments to the protocol, funding, recruitment, personnel or resource input required of NHS Lothian.

Substantial amendments to the protocol will require approval from the ethics committee which approved your study and the MHRA where applicable.

Please keep this office informed of the following study information:

1. Date you are ready to begin recruitment, date of the recruitment of the first participant and the quarterly recruitment figures thereafter.
2. Date the final participant is recruited and the final recruitment figures.
3. Date your study / trial is completed within NHS Lothian.

I wish you every success with your study.

Yours sincerely

Fiona McArdle.

Ms Fiona McArdle
Deputy R&D Director

cc: Ms Caroline Brydon, Radiology, RIE
Ms Catriona Rostron, Associate Nurse Director, NHS Lothian
Ms Dawn Cardy, Imaging Study Information Administrator, CRIC, QMRI
Ms Dawn Lyster, Labs, NHS Lothian
Ms Shuna Colville, Research Nurse, RIE

Appendix 8: Participant Information Sheet for study in Chapter 4 & 5

Participant Information Sheet: patients Movement

fatigue in multiple sclerosis

We would like to invite you to take part in our research study. Before you decide we would like you to understand why the research is being done and what it would involve. One of our team will go through the information sheet with you and answer any questions you have. This should take about 20 minutes. Ask us if there is anything that is not clear.

This study involves individuals with multiple sclerosis, who exercise regularly (at least for 30 minutes twice weekly and can walk/jog for a mile) with no or a slight neurological disability. This study does not involve testing drugs.

What is the purpose of the study?

The aim of this pilot study is to measure the symptoms noticed during and immediately following physical activity in a laboratory setting at regular intervals over a 2 year period, and metabolic changes in the brain using scans.

Why have I been invited?

You have been selected because you have MS and exercise regularly. During exercise, you may have noticed exercise related symptoms such as leg weakness, foot drop and difficulty controlling your leg(s) that are temporary and recover with rest.

Do I have to take part?

Taking part in this study is entirely voluntary. If you agree to take part then we will ask you to sign a consent form. You can change your mind at any point during the study and withdraw without having to give a reason. If you withdraw, the standard

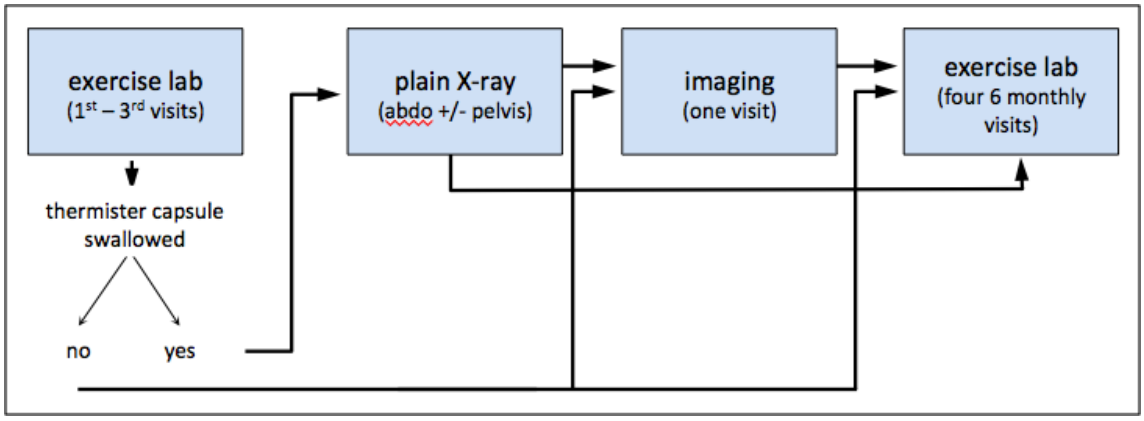
of your care will not be affected in any way.

What will happen to me if I take part?

You are invited to visit Anne Rowling Clinic for an interview (approximately 30 minutes) and consent. You will then be invited to the exercise laboratory at Queen Margaret University, for maximum of seven visits, each lasting **a maximum of** 2 hours.

As an optional step before your second visit to Queen Margaret University, you may swallow a capsule to record your core body temperature before, during and after exercise. The capsule is for single use and it has not been administered before. The capsule will be passed in your stools usually within 2-3 days. If not passed within this time, you may take a laxative to facilitate its passage. You are not expected to retrieve the capsule from your stools.

You may be invited to have brain scans on one occasion, lasting no more than 60 minutes. If you have taken the capsule, we have to perform a plain X-ray of your abdomen, and if necessary pelvis, before the scan (anticipated to last 5 minutes) to make sure that the capsule is no longer in your body. We may advise you to have a blood test (one teaspoon of blood removed from one arm using standard methods) to check your kidney function.



Expenses and payments

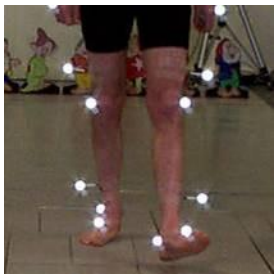
This study involves traveling to Anne Rowling Clinic (Little France, Edinburgh), Queen Margaret University (Musselburgh) and Clinical Research Imaging Centre (Little France, Edinburgh). Travel expenses from your home or work place, including car-parking fees, will be reimbursed upon presentation of the receipts. Please note that there are no inducements or additional payments for taking part in this study.

What will I have to do?

The **interview** with a neurologist will take place at Anne Rowling Clinic (ARC) and he/she will go through this information sheet, answer any questions and provide the consent form. You have up to 1 month to return the consent form, if you wish to take part.

You will then be sent, by post or email, dates to attend the exercise laboratory at Queen Margaret University (QMU) by the researchers and also two fatigue questionnaires for you to complete and return by stamped addressed envelope to ARC or email. Please bring comfortable clothes and trainers (running shoes) to exercise.

Visit 1 to Queen Margaret University.



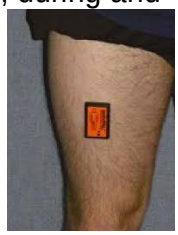
The first visit to the exercise laboratory is to get accustomed to the set up while the researchers develop an individualized exercise program based on your usual exercise regime. Small balls covered in reflective material will be attached on to your pelvis and lower limbs using skin friendly double sided tape. You are not expected to exercise outside your comfort zone. You are then given a second appointment to attend QMU to **record your walking/running pattern before, during and after a 20-minute exercise task** and any exercise related symptoms.

Visit 2 to Queen Margaret University.

Before your second visit you may swallow a capsule **(optional)**, containing a sensor (optional), to measure your core body temperature before, during and after **a 20-minute** exercise task. It does not contain any medication. The size of the capsule is shown relative a 20 pence coin. Further details of the capsule can be found on manufacturer's website (**www.equivital.co.uk**). Researcher will provide the capsule. Please swallow it before your appointment time and avoid taking any heavy meals with it or after you have taken it.

You can take drinks kept at room temperature. Please avoid hot or cold drinks from swallowing the capsule to 3 hours after completing exercise. The capsule is usually passed in your stools within 2-3 days. We will advise you to have a plain X-ray (abdominal) 1 week after taking the capsule, to ensure that it has passed.

The assessment of your walking/running pattern will take place before, during and after the exercise task. During this task, you may jog/run at a self-selected speed for 20 minutes, with 3 minutes for warm-up and 2 minutes to adjust to your usual speed. **You will receive an activity monitor (small device) that needs to be attached to your thigh with double sided tape for seven days, and which will record the number of steps you take**.



Visit 3 to Queen Margaret University.

The third visit (within 4 weeks from the second) is to determine whether the recordings are reproducible. The exercise tasks and gait assessments will be the same as in visit 2, but without having to swallow the capsule.

Visit to Clinical Research Imaging Centre for brain imaging.

You are given the option to have brain scans using magnetic resonance imaging. Scanning time will be no more than 60 minute in one visit to the Clinical Research Imaging Centre (CRIC), Royal Infirmary of Edinburgh. You must NOT enter the MRI scanner until you know that the capsule is no longer in your body. This is because it contains metal and can cause internal damage to your body when in the scanner.
The plain X-ray(s) will assess safety.

We use foam pads to place your head comfortably in the correct position. During the scan you are asked to gently move your foot, hand or mouth at different times (for less than 2 minutes) while in the scanner. You may receive a contrast agent, called gadolinium, after placing a cannula into a vein on the back of your hand or front of the elbow.

Subsequent visits to the exercise laboratory at Queen Margaret University.

Once you have completed the third visit to the exercise laboratory you are invited for four more visits, at six monthly intervals. The exercise regime will be the same as before. Prior to each visit, the same fatigue questionnaires will be sent to you for completion and return by mail or email. If you do not feel up to exercising on a given date and would like to rearrange the appointment then please get in touch with Dr Don Mahad or Dr Marietta van der Linden (see below).

What are the possible risks and disadvantages of taking part?

When exercising in an unfamiliar setting there is a potential risk of injury. We will

minimize this by making sure that you are comfortable with the set up at each visit. There is a low chance of unexpected complications such as chest pain occurring during exercise. If this occurs, we will arrange for a medical assessment. The researcher may notice changes in your exercise pattern on the recording. If you wish to know of any changes please contact the Chief Investigator (Don Mahad). We recommend that you do not exert beyond your comfort zone for a couple of days before each visit.

To secure the markers to your pelvis and legs for the assessment of your gait, we use **double sided 'toupee'** skin friendly tape, which can cause redness of the skin after the tape is removed. If you have an allergy please inform the researchers. The capsule to record core body temperature has been shown to be safe in MS patients and healthy individuals. The researchers will check that you do not have a condition, which will put you at risk from the capsule. MRI of any part of your body should not be performed until it is established with certainty that the capsule is no longer in your body.

As part of the MRI you may be advised to have a dye or contrast agent (gadolinium) injected using a cannula placed in a vein on the back or your hand or front of the elbow. If you have kidney disease or take Metformin please alert the researchers. This is because of the risk of developing a kidney scaring disease.

What are the side effects of any treatment received when taking part?

The study is not a clinical trial and you will not be expected to take a drug. The plain X-rays of the abdomen, and if necessary the pelvis, are performed approximately 1 week after swallowing the capsule to ensure it has passed from the bowel. The amount of radiation exposure by these X-rays is comparable to 7 months of natural background radiation in the UK population. There is a small risk of developing a fatal cancer as a result of exposure to this amount of radiation and the risk is approximately 1 in 20,000.

What are the possible benefits of taking part?

The information we will get from this study will help us to develop methods to measure MS symptoms during exercise and determine they may changes over time.

What happens when the research study stops?

We will analyse the data from all subjects within 6 months from completion of the study and let you know the outcome, if you wish to be informed.

What if there is a problem?

Any complaint about the way you have been dealt with or any possible harm you might suffer will be addressed. The detailed information on this is given below.

Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and handle information confidentially .

What if relevant new information becomes available?

Exercise recordings and brain scans may show changes over time. If you wish to know, the research doctor will inform you and explain any changes. Any change in your recordings is unlikely to influence clinical management, as this is an experimental study.

What will happen if I don't want to carry on with the study?

You are free to withdraw at any point and you do not have to provide a reason. We will use your anonymised data collected until the time of your withdrawal.

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researchers or Shuna Colville [shuna.colville@ed.ac.uk], who is the clinic manager for the Anne Rowling Regenerative Neurology Clinic. They will try to answer your questions [donmahad@nhs.net]. If you remain unhappy and wish to complain formally, please contact NHS Lothian:

NHS Lothian Complaints Team
2nd Floor
Waverley Gate
2-4 Waterloo Place
Edinburgh
EH1 3EG
Tel: 0131 465 5708
Email: craft@nhslothian.scot.nhs.uk

If there is harm resulting from the study, as a result of any tests or procedures

you received, appropriate compensation will be available and details of indemnity scheme will be provided to you on request. We would not be bound to pay compensation where the injury resulted from a drug or procedure outside the trial protocol and when the protocol was not followed.

Will my taking part in the study be kept confidential?

All information, collected about you during the course of the research, will be kept strictly confidential. The recordings of your movements during exercise will only be stored in the form of a cartoon that joins the detectors placed on your body. These cartoons will not contain identifiable images of you. Any information about you which leaves the department will have your name and address removed so that you cannot be identified. Your data will be collected using paper and electronic medical records by researchers and stored under a unique number. An electronic file containing your CHI number to match with the unique number and identifiable data will be kept in a password protected NHS computer held in a secure room at Anne Rowling Clinic until the end of the study. The chief investigator (Don Mahad) will have access to this file. In addition, regulatory authorities and research and development audit committee may access your records, only for monitoring of the quality of the research. This anonymised code will be used in all subsequent dealing with data analysis and interpretation. The identifiable data will be deleted without any other form of data storage once the study is completed (anticipated to be two and a half years from the start date of the study). If your data needs to be retained for future studies the researcher will request approval from the local ethics committee.

Involvement of your family doctor

If you decide to take part in this study, we recommend that a letter is sent to your GP to explain the nature of the study and possible implications for your clinical care.

What will happen to the results of the study?

Results of this pilot study will be used to design future studies, as this is a relatively small study. If you wish to be informed of the results we will do so in writing and include your results as well as a summary of the overall findings.

Who is organizing and funding the research?

This research is organized by the University of Edinburgh and funded by Edinburgh Bioquarter.

Who has reviewed the study?

This study was review by an independent group of people, Research Ethics Committee, to protect your interests. The scientific aspects have been reviewed by Edinburgh Bioquarter funding committee.

Further information and contact details

General information about the researcher and their scientific background can be accessed online at the following websites:
Don Mahad: http://www.cnr.ed.ac.uk/Research/mahad.html (email: don.mahad@ed.ac.uk)
Tom Mercer: http://www.qmu.ac.uk/ph/staff_biogs/tom_mercer.htm (email: TMercer@qmu.ac.uk
Marietta van der Linden: http://www.qmu.ac.uk/ph/staff_biogs/mariette_van_der_linden.htm (email: MVanDerLinden@qmu.ac.uk)

Specific information about the thermistor capsule can be obtained from the manufacturer (www.ksi.uconn.edu). For specific information about the study and advice as to whether you should participate, please contact Don Mahad.

If you are unhappy and would like to discuss please contact Don Mahad. If you wish to submit an official complaint please visit the following website.
http://www.nhslothian.scot.nhs.uk/YourRights/ComplimentsConcernsComplaints/Pages/MakingAComplaint.aspx

Appendix 9: Informed consent form for study in Chapter 4 & 5

CONSENT FORM: patients

Title of project: Movement fatigue in multiple sclerosis

Name of Researcher: Dr Don Mahad

Please initial box

1. I confirm that I have read and understand the information sheet dated April 2015 (version 3) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐

3. "I understand that relevant sections of [my medical notes and data collected during the study may be looked at by individuals from the regulator authorities and from the Sponsor(s) (NHS Lothian and the University of Edinburgh) or from the/other NHS Board(s) where it is relevant to my taking part in this research. I give permission for those individuals to have access to my records" ☐

4. I agree to take the thermistor capsule, have X-ray(s) and understand the issues relating to radiation and MRI ☐

5. I agree to wear the activpal device for a minimum of 5 days out of 10 following exercise ☐

6. I agree to undergo MRI scans ☐

7. I agree to my GP being informed of my participation in the study ☐

8. I agree to take part in the above study. ☐

Name of Patient	Date	Signature

Name of Person taking	Date Signature consent

When completed: 1 for participant; 1 for researcher site file; 1 (original) to be kept in medical notes.

Appendix 10: Instructions for ActivPal (Chapter 5)

Thank you for wearing the activity monitor. It records how much of the day you are sitting, standing or walking.
We would like you to wear it for at least five days, you can take the monitor off your leg during the night, but this means attaching it to your thigh in the morning as soon as possible after you get up, and taking it off before you go to bed.
The monitor only records for 10 days starting from the time you were given the monitor so please could you wear it next day of receiving it!
After you have worn the monitor for five days, could you please return it in your next visit. We can then download the data from the monitor onto a computer and see how much of the day you spent either walking, standing or lying/sitting.

Applying the activPAL (activity monitor)
It is most comfortable to wear the monitor attached to the thigh (see picture). It should be positioned on the midline of the thigh, between the hip and the knee. However, it will function correctly if placed anywhere on the front of the thigh in the orientation indicated by the figure on the front panel. (the little man should be standing/UP pointing upwards)
The monitoring device and the tape are not waterproof and they should be removed for bathing. The skin should be thoroughly dried after bathing to maximise the adherence of the tape/gel.

Finally, please return the included form and note which days the monitor was attached to your thigh.

Please contact me on 07873854432 if you have any questions regarding the monitor.

Thank you very much for helping with this study,

Georgia Andreopoulou
PhD student
Physiotherapy Division
School of Health Sciences
Queen Margaret University Drive
Musselburgh
EH21 6UU
Email: GAndreopoulou@qmu.ac.uk

	Days & dates	time
Day 1		
Day 2		
Day 3		
Day 4		
Day 5		

Comments

_____
_____
_____
_____
_____
_____
_____
_____

Please fill in the day and the times above

Appendix 11: Borg Scale of Perceived Exertion (Borg, 1982)

BORG RATING OF PERCEIVED EXERTION (RPE) SCALE

Number rating	Verbal rating
6	
7	Very, very light
8	
9	Very light
10	
11	Fairly light
12	

13	Somewhat hard
14	
15	Hard
16	
17	Very hard
18	
19	Very, very hard
20	

Appendix 12: Feeling Scale (Hardy & Rejeski, 1989)

While participating in exercise, it is common to experience changes in mood. Some individuals find exercise pleasurable, whereas others find it to be unpleasant. Additionally, feeling may fluctuate across time. That is, one might feel good and bad a number of times during exercise. Scientists have developed this scale to measure such responses.

+5 Very good

+4

+3 Good

+2

+1 Fairly good

0 Neutral

-1 Fairly bad

-2

-3 Bad

-4

-5 Very bad

Appendix 13: Fatigue Scale for Motor and Cognitive Functions (Penner et al, 2009)

	Does not apply at all	Does not apply much	Slightly applies	Applies a lot	Applies completely
1. When I concentrate for a long time, I get exhausted sooner than other people of my age.	◯	◯	◯	◯	◯
2. When I am experiencing episodes of exhaustion, my movements become noticeably clumsier and less coordinated.	◯	◯	◯	◯	◯
3. Because of my episodes of exhaustion, I now need more frequent and/or longer rests during physical activity than I used to.	◯	◯	◯	◯	◯
4. When I am experiencing episodes of exhaustion, I am incapable of making decisions.	◯	◯	◯	◯	◯
5. When faced with stressful situations, I now find that I get physically exhausted quicker than I used to.	◯	◯	◯	◯	◯
6. Because of my episodes of exhaustion, I now have considerably less social contact than I used to.	◯	◯	◯	◯	◯
7. Because of my episodes of exhaustion, I now find it more difficult to learn new things than I used to.	◯	◯	◯	◯	◯
8. The demands of my work exhaust me mentally more quickly than they used to.	◯	◯	◯	◯	◯
9. I feel the episodes of exhaustion particularly strongly in my muscles.	◯	◯	◯	◯	◯
10. I no longer have the stamina for long periods of physical activity that I used to have.	◯	◯	◯	◯	◯

11. My powers of concentration decrease considerably when I'm under stress.	◯	◯	◯	◯	◯
12. When I am experiencing episodes of exhaustion, I am less motivated than others to start activities that involve physical effort.	◯	◯	◯	◯	◯
13. My thinking gets increasingly slow when it is hot.	◯	◯	◯	◯	◯
14. When I am experiencing an episode of exhaustion, my movements become noticeably slower.	◯	◯	◯	◯	◯
15. Because of my episodes of exhaustion, I now feel less like doing things which require concentration.	◯	◯	◯	◯	◯
16. When an episode of exhaustion comes on, I am simply no longer able to react quickly.	◯	◯	◯	◯	◯
17. When I am experiencing episodes of exhaustion, certain words simply escape me.	◯	◯	◯	◯	◯
18. When I am experiencing episodes of exhaustion, I lose concentration considerably quicker than I used to.	◯	◯	◯	◯	◯
19. When it is hot, my main feeling is one of extreme physical weakness and lack of energy.	◯	◯	◯	◯	◯
20. During episodes of exhaustion, I am noticeably more forgetful.	◯	◯	◯	◯	◯

Appendix 14: Fatigue Severity Scale (Krupp et al, 1989)

Please circle the number between 1 and 7 which you feel best fits the following statements. This refers to your usual way of life within the last week. 1 indicates "strongly disagree" and 7 indicates "strongly agree."

Read and circle a number.	Strongly Disagree → Strongly Agree						
1. My motivation is lower when I am fatigued.	1	2	3	4	5	6	7
2. Exercise brings on my fatigue.	1	2	3	4	5	6	7
3. I am easily fatigued.	1	2	3	4	5	6	7
4. Fatigue interferes with my physical functioning.	1	2	3	4	5	6	7
5. Fatigue causes frequent problems for me.	1	2	3	4	5	6	7
6. My fatigue prevents sustained physical functioning.	1	2	3	4	5	6	7
7. Fatigue interferes with carrying out certain duties and responsibilities.	1	2	3	4	5	6	7
8. Fatigue is among my most disabling symptoms.	1	2	3	4	5	6	7
9. Fatigue interferes with my work, family, or social life.	1	2	3	4	5	6	7

Appendix 15: Ethical opinion from NHS Research Ethics Committee (Chapter 6)

NHS
Health Research Authority

West Midlands - Edgbaston Research Ethics Committee

The Old Chapel
Royal Standard Place
Nottingham
NG1 6FS

24 July 2018

Georgia Andreopoulou
School of Health Sciences
Queen Margaret University Drive
Musselburgh
EH21 6UU

Dear Ms Andreopoulou

Study title:	**The direct orthotic effect of functional electrical stimulation on gait kinematics and walking speed in people with MS under dual-tasking and fatiguing walking conditions simulating daily life.**
REC reference:	**18/WM/0062**
Protocol number:	**NA**
Amendment number:	**1**
Amendment date:	**27 June 2018**
IRAS project ID:	**240383**

The above amendment was reviewed on 23 July 2018 by the Sub-Committee in correspondence.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

The Sub-Committee agreed that the substantial amendment did not raise any material ethical issues.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Non-validated questionnaire [Foot drop questionnaire]	1	27 June 2018
Notice of Substantial Amendment (non-CTIMP)	1	27 June 2018
Other [Letter of amendment]		
Participant information sheet (PIS) [Healthy PIS Tracked]	3	27 June 2018
Participant information sheet (PIS) [PIS for patients]	3	27 June 2018
Participant information sheet (PIS) [Healthy PIS Clean]	3	27 June 2018

Appendix 16: NHS Lothian Research and Development approval (Chapter 6)

University Hospitals Division

NHS Lothian

Queen's Medical Research Institute
47 Little France Crescent, Edinburgh, EH16 4TJ

FM/CK/approval

10 April 2018

Research & Development
Room E1.16
Tel: 0131 242 3330

Email:
accord@nhslothian.scot.nhs.uk

Director: Professor Tim Walsh

Miss Georgia Andreopoulou
Queen Margaret University
Edinburgh

Dear Miss Andreopoulou

Lothian R&D Project No: 2018/0074 **REC No:** 18/WM/0062

Title of Research: The direct orthotic effect of functional electrical stimulation on gait kinematics and walking speed in people with MS under dual tasking and fatiguing walking conditions simulating daily life

Participant Information Sheet:
(Patient) version 2.0, dated 6 March 2018
(Healthy) version 2.0, dated 6 March 2018

Consent Form:
(Patient) version 3.0, dated 26 March 2018
(Healthy) version 2.0, dated 26 March 2018

Protocol: version 1.0, dated 12 January 2018

I am pleased to inform you this letter provides Site Specific approval for NHS Lothian for the above study and you may proceed with your research, subject to the conditions below.

We note that NHS Lothian is participating in this trial as a Participant Identification Centre (PIC).

Please note that the NHS Lothian R&D Office must be informed of any changes to the study such as amendments to the protocol, funding, recruitment, personnel or resource input required of NHS Lothian.

Substantial amendments to the protocol will require approval from the ethics committee which approved your study and the MHRA where applicable.

Please keep this office informed of the following study information, **which is a condition of NHS Lothian R&D Management Approval**:

1. Date you are ready to begin recruitment, date of the recruitment of the first participant and the monthly recruitment figures thereafter.
2. Date the final participant is recruited and the final recruitment figures.
3. Date your study / trial is completed within NHS Lothian.

I wish you every success with your study.

Yours sincerely
Fiona McArdle

Ms Fiona McArdle
Deputy R&D Director

Cc: Shuna Colville, Anne Rowling Clinic, University of Edinburgh
Michael Pearson, DCN, NHS Lothian

STA034

Appendix 17: Participant Information Sheet for study in Chapter 6

Queen Margaret University

EDINBURGH

Participant information sheet

Study title: The effects of FES in a variety of walking conditions in people with MS

We would like to invite you to take part in our research study. Before you decide we would like you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you like. Ask us if there is anything not clear or you would like more information.

What is the purpose of the study?

The aim of this study is to measure the benefits of functional electrical stimulation (FES) to treat foot drop in 'real life' conditions, i.e. walking while doing another

task that requires your attention and when tired. Previous studies have investigated the benefits of FES, but only in people who were just walking up and down a laboratory, which may not be an adequate reflection of the benefits people get from using FES in daily life. This study therefore will look at the benefits of FES under conditions that are closer to those in 'real life'.

Why have I been invited?

You have been invited because you have been diagnosed with Multiple Sclerosis, you are experiencing foot drop while walking and you are currently using FES.

Do I have to take part?

No, it is entirely up to you whether or not you take part in this study. Your clinical care will not be affected by your decision. You can change your mind and withdraw from the study at any point, without giving any reason. If you decide to take part in the study, you will be asked to sign an informed consent form. If you do withdraw, any data that have been collected will be retained.

What will happen if I take part?

If you are interested in taking part in this study, you will be invited to Queen Margaret University (QMU) for two visits. In the first visit, the nature of the study will be explained to you. After clarifying any further questions that you may have,

Figure 1 Stroop test

and you agree to take part in the study, you will be asked to sign an informed consent form. This first visit will last approximately one hour and will involve measuring your weight, height and leg length.. We will also explain the so-called Stroop test which we will ask you to do in the second visit. In this test, the words of four colours, but written with a different colour, will be projected on the wall in front of you and you will be asked to identify the colour of the text and not the word itself (see fig.1). You will be given a few practice trials.

The second visit will last approximately two and a half hours and you will be asked to do eight walks (7 meters each) under the three conditions (24 walks in total). Condition A: Normal walking in a straight line with FES on (four walks) and off (four walks). Condition B: Walking whilst performing the Stroop test (explained above) with FES on (four walks) and off (four walks). Condition C: This walking condition is the same as B, but prior to this condition we will ask you to do a 'fatiguing task' i.e. a task that aims to make you feel (physically) tired. For the fatiguing task, we will ask you to do an 'incremental shuttle walk test' in which involves walking between two cones (10m distance) several times. The time you have to cover the distance between the two cones (signifies by 'bleeps') will become increasingly shorter, until you cannot get to the next cone before the next 'bleep'. This test will last a maximum of 20 minutes.

During the eight walks in each of the three conditions, the movement of your legs and pelvis will be recorded using special cameras that capture the movement of the markers. Markers are small balls covered by a reflective material, which will be placed onto your legs and pelvis using skin-friendly tape. The markers and their positioning are shown in Figure 2. You can stop at any point between the walks described above. In order to make sure you feel safe in the laboratory environment a person will walk next to you.

Figure 2 Marker placement

What are the possible disadvantages and risks of taking part?

We believe that there are no serious risks of taking part. The walking conditions are simulating those in daily life, such as walking whilst doing a mental task or walking when tired. If you do experience pain or discomfort, you can stop at any time. If there is an unexpected complication, the University staff are certified first aiders. There is a first aid kit at the laboratory and health and safety assessments have been carried out to ensure that the laboratory environment is safe and clear of obstacles which can cause trips or falls. We will make sure that it is clear what you will be asked to do in each of the conditions.

What are the possible benefits of taking part?

While the benefits of the study are not directly linked with you, the results will inform us regarding the benefit of FES while walking and doing a task that requires your attention and when physically tired.

Expenses and payments

The study involves travelling to Queen Margaret University (Musselburgh) twice. Upon presentation of receipts, travel expenses up to £20 will be reimbursed. There are no payments or inducements for taking part in this study.

Involvement of your family doctor

If you decide to take part in the study, your GP will be notified by a letter explaining the nature of the trial. You could also make other healthcare professionals that are involved in your treatment aware of your participation in the trial.

What happens when the research study stops?

The data from all the participants will be analysed upon completion of the study and we will let you know about your personal and general outcomes if you wish.

Will my taking part in the study be kept confidential?

Yes. All information collected about you during the study will be kept strictly confidential. Every participant is given a unique code at the beginning of the study and only this code is associated with any recorded data. The data recorded on paper, such as questionnaires, will be stored in locked cabinets to which only the researcher will have access. You will not be identifiable from the electronic data that we record with the cameras (fig. 3) and these data will be stored on a password-protected network drive at Queen Margaret University. Only the researchers involved in the study will have access to your data and on a few occasions, the research committee responsible for monitoring the quality of the research. All data will be stored for at least five years. We will make sure that all identifiable information will be removed from any presented or published work produced from the data collected.

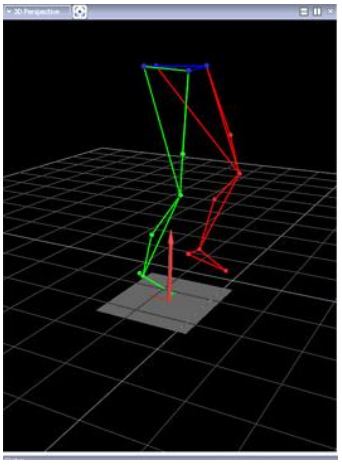


Figure 3 Electronic recording from the cameras.

What will happen to the results of the study?

The results of this study will be submitted for publications in scientific journals and/or conferences. We will make sure that you will not be identifiable in any of the manuscripts published. If you wish to be informed of your personal results, we will do so by sending you a summary of your findings.

What if there is a problem?

Any complaint about the study, the way you have been dealt with during the study or if you believe you have been harmed in any way by taking part will be addressed. The university has a liability insurance scheme for compensation in case of harm caused by negligence of the researcher, but no compensation for any injury resulting from a drug or any procedure outside the abovementioned protocol.

If you have any concerns or complaints throughout the study please contact Kim Stuart, Head of Research and Knowledge Exchange Development Unit at Queen Margaret University during office hours.

Kim Stuart

Head of Research and Knowledge Exchange Development Unit

Queen Margaret University

Queen Margaret University Drive

Musselburgh

East Lothian

EH21 6UU

Tel: 01314740000 (ask for Kim Stuart when prompted)

E-mail: kstuart@qmu.ac.uk

Who is organizing and funding the study?

The study is funded by a PhD bursary awarded by Queen Margaret University. It is organized by a university research team that will not receive any additional income for being involved in the study.

Who has reviewed the study?

This study was given a favourable ethical opinion by the Edgbaston-West Midlands Research Ethics Committee and the NHS Lothian Research and Development (R&D) office. The Queen Margaret University Research Ethics Committee also reviewed the study.

Further information and contact details

Thank you for your valuable time. If you need any further information, please contact us.

Georgia Andreopoulou, MSc
PhD candidate
Queen Margaret University
Queen Margaret University Drive
Musselburgh
EH21 6UU

Dr Marietta van der Linden
Senior Research Fellow (PhD supervisor)
Queen Margaret University
Queen Margaret University Drive
Musselburgh
EH21 6UU

Tel: 07873854432

gandreopoulou@qmu.ac.uk

Tel: 0131 474 0000 (ask for Marietta van der Linden when prompted)

mvanderlinden@qmu.ac.uk

If you need an independent advice, please contact:

Gill X. Murray

Advanced Physiotherapy Practitioner in Neurology

Physiotherapy Department

Slateford Medical Centre

27 Gorgie Park Close

Edinburgh, EH14 1NQ

0131 455 9850 mobile 07816 174 293

(Monday, Wednesday, Thursday, Friday)

Gill.X.Murray@nhslothian.scot.nhs.uk

Appendix 18: Informed consent form for study in Chapter 6

Queen Margaret University

EDINBURGH

CONSENT FORM

Title of Project: **The effects of FES in a variety of walking conditions in people with MS**

Name of Researcher: Georgia Andreopoulou

1. I confirm that I have read the information sheet dated....26/3/2018..... (version 3) for the
   above study. I have had the opportunity to consider the information, ask questions and have
   had these answered satisfactorily.

2. I understand that my participation is voluntary and that I am free to withdraw at any time
   without giving any reason, without my medical care or legal rights being affected.

3. I agree to my General Practitioner being informed of my participation in the study.

4. I understand that during my involvement in this study, the PI will have access to my personal information (i.e. home address and email) to arrange my study visits to the university.

5. I understand that the Sponsor of the study might review my medical records.

6. I agree to take part in the above study.

_____	_____	_____
Name of Participant	Date	Signature
_____	_____	_____
Name of Person taking consent	Date	Signature

Appendix 19: Advertisement for healthy volunteers

Queen Margaret University

EDINBURGH

Volunteers please!

- ✓ Are you between the ages of 18-80?
- ✓ Are you free of any musculoskeletal injuries or neurological disorders?
- ✓ Are you interested in learning more about your walking?

If yes, please read on:
We at Queen Margaret University (Musselburgh) are running a study to assess the walking patterns and walking speed of people with and without multiple sclerosis. Volunteer participants will be asked to attend QMU campus for one single visit which would last a maximum of two and a half hours. In this visit, we will record your walking pattern while you are walking and performing an attention-demanding task at the same time and repeat this again after a physically demanding walking task. Travel costs to and from the University up to £20 will be reimbursed.

This study has received a favourable opinion from the West Midlands NHS Research Ethics Committee and Queen Margaret University Research Ethics Committee.
If you interested in taking part in this study and would like more information, please contact Georgia Andreopoulou at Queen Margaret University (07873854432 or gandreopoulou@qmu.ac.uk)

Appendix 20: Multiple Sclerosis Walking Scale

If you cannot walk at all, please tick this box ☐

In the past two weeks, how much has your MS...	Not at all	A little	Moderately	Quite a lot	Extremely
1. Limited your ability to walk?	1	2	3	4	5
2. Limited your ability to run?	1	2	3	4	5
3. Limited your ability to climb up and down stairs?	1	2	3	4	5
4. Made standing when doing things more difficult?	1	2	3	4	5
5. Limited your balance when standing or walking?	1	2	3	4	5
6. Limited how far you are able to walk?	1	2	3	4	5
7. Increased the effort needed for you to walk?	1	2	3	4	5
8. Made it necessary for you to use support when walking indoors (eg holding on to furniture, using a stick, etc.)?	1	2	3	4	5
9. Made it necessary for you to use support when walking outdoors (eg using a stick, a frame, etc.)?	1	2	3	4	5
10. Slowed down your walking?	1	2	3	4	5
11. Affected how smoothly you walk?	1	2	3	4	5
12. Made you concentrate on your walking?	1	2	3	4	5

From the numbers you circle against these questions, your healthcare professional can calculate your MSWS-12 score. This is done by adding the numbers you have circled, giving a total out of 60, and then transforming this to a scale with a range from 0 to 100. Higher scores indicate a greater impact on walking than lower scores.

Appendix 21: Multiple Sclerosis Neuropsychological Screening Questionnaire (Benedict et al, 2003)

Name:

Date:

Sex (circle one): Male / Female

INSTRUCTIONS: The following questions ask about problems that you may experience. Rate how often these problems occur **AND** how severe they are. Base your ratings on how you have been over the **last 3 months.**
Please check the appropriate box.

1. Are you easily distracted?	r	r	r	r
2. Do you lose your thoughts while listening to somebody speak?	r	r	r	r
3. Are you slow when trying to solve problems?	r	r	r	r
4. Do you forget appointments?	r	r	r	r
5. Do you forget what you read?	r	r	r	r
6. Do you have trouble describing shows or programs recently watched?	r	r	r	r
7. Do you need to have instructions repeated?	r	r	r	r
8. Do you have to be reminded to do tasks?	r	r	r	r
9. Do you forget errands that were planned?	r	r	r	r
10. Do you have difficulty answering questions?	r	r	r	r
11. Do you have difficulty keeping track of two things at once?	r	r	r	r
12. Do you miss the point of what someone is trying to say?	r	r	r	r
13. Do you have difficulty controlling impulses?	r	r	r	r
14. Do you laugh or cry with little cause?	r	r	r	r
15. Do you talk excessively or focus too much on your own interests?	r	r	r	r

Appendix 22: Hospital Anxiety and Depression Scale

Tick the box beside the reply that is closest to how you have been feeling in the past week.

Don't take too long over you replies: your immediate is best.

D	A		D	A	
		I feel tense or 'wound up':			**I feel as if I am slowed down:**
	3	Most of the time	3		Nearly all the time
	2	A lot of the time	2		Very often
	1	From time to time, occasionally	1		Sometimes
	0	Not at all	0		Not at all
		I still enjoy the things I used to enjoy:			**I get a sort of frightened feeling like 'butterflies' in the stomach:**
0		Definitely as much		0	Not at all
1		Not quite so much		1	Occasionally
2		Only a little		2	Quite Often
3		Hardly at all		3	Very Often
		I get a sort of frightened feeling as if something awful is about to happen:			**I have lost interest in my appearance:**
	3	Very definitely and quite badly	3		Definitely
	2	Yes, but not too badly	2		I don't take as much care as I should
	1	A little, but it doesn't worry me	1		I may not take quite as much care
	0	Not at all	0		I take just as much care as ever
		I can laugh and see the funny side of things:			**I feel restless as I have to be on the move:**
0		As much as I always could		3	Very much indeed
1		Not quite so much now		2	Quite a lot
2		Definitely not so much now		1	Not very much
3		Not at all		0	Not at all
		Worrying thoughts go through my mind:			**I look forward with enjoyment to things:**
	3	A great deal of the time	0		As much as I ever did
	2	A lot of the time	1		Rather less than I used to
	1	From time to time, but not too often	2		Definitely less than I used to
	0	Only occasionally	3		Hardly at all
		I feel cheerful:			**I get sudden feelings of panic:**
3		Not at all		3	Very often indeed
2		Not often		2	Quite often
1		Sometimes		1	Not very often
0		Most of the time		0	Not at all
		I can sit at ease and feel relaxed:			**I can enjoy a good book or radio or TV program:**
	0	Definitely	0		Often
	1	Usually	1		Sometimes
	2	Not Often	2		Not often
	3	Not at all	3		Very seldom