2015

THE EFFECTS OF PILATES BASED CORE STABILITY TRAINING IN PEOPLE WITH MS

Esther Elizabeth Fox

PHD

THE EFFECTS OF PILATES BASED CORE STABILITY TRAINING IN PEOPLE WITH MS

Fox, Esther Elizabeth

Award date:
2015

Awarding institution:
University of Plymouth

THE EFFECTS OF PILATES BASED CORE STABILITY TRAINING IN PEOPLE WITH MS

BY

ESTHER FOX

A thesis submitted to Plymouth University in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of Health Professions

Faculty of Health and Human Sciences

Abstract

Background: People with Multiple Sclerosis experience difficulties with balance and mobility. Pilates exercises are often used to address these difficulties.

Design: This was a multi-centre, double blind, block randomised, controlled trial. Eligible participants were recruited from seven UK centres. Participants were randomly allocated to either: Pilates based core stability training (Pilates), Standardised Exercise (SE) or Relaxation (placebo). All received face-to-face training sessions over a 12 week period; together with a home exercise programme. Blinded assessments were taken before training, at the end of the 12 week programme and at 16 weeks (follow-up).

The primary outcome measure was the 10metre timed walk (10mtw). Secondary outcome measures were the MS walking Scale (MSWS-12), Functional Reach Test (FRT) (forwards and lateral), a 10 point Visual Analogue Scale (VAS) to determine “Difficulty in carrying a drink when walking”, and the Activities-specific Balance Confidence (ABC) Scale. Effects on deep abdominal muscles were measured with ultrasound imaging (USI) in a subgroup of patients.

Independent t-tests were performed to compare groups. Sensitivity analyses were undertaken to confirm the results. A mixed factorial ANOVA analysed the effect of intervention over time upon TrAb and IO upon USI.

Results: Of the 100 participants recruited, 13 relapsed leaving 94 for intention to treat analysis. At 12 weeks there were significant differences between:

(1) Pilates and Relaxation for walking velocity ($p=0.04$), forward ($p=0.04$) and lateral ($p=0.04$) FRT.

(2) SE and Relaxation for all measures ($p<0.05$) apart from the VAS. These remained at 16 weeks for 10mtw ($p=0.04$), LFR ($p<0.01$) MSWS-12 ($p=0.03$) and ABC ($p= 0.03$).

There were no significant interactions ($p>0.05$) between groups or over time for TrAb and IO.

Conclusions: Participants improved with both Pilates and SE in the short term; with broader and longer-lasting effects in the SE group. USI did not detect any effect of group over time.

Table of Contents

List of figures

List of Tables

Tables in Appendices

Acknowledgements

I would like to give my absolute heartfelt thanks and admiration to Dr Jenny Freeman as my director of studies. Without her none of this PhD would have been possible. Her endless guidance, patience, support and belief in me has been what can only be described as invaluable and far beyond the call of duty.

In addition, heartfelt thanks go to Dr Alan Hough, my second supervisor, who believed in me from the beginning of my life as a Physiotherapist in 2005 and who has been an excellent mentor since in supporting me through my clinical and academic career. One could not wish for better PhD supervisors.

I would like to thank also my family, especially Mum and John, not to mention 'Strict' Auntie Carol for keeping me on the straight and narrow. And to Sophie, Gail, Beth and Jenny for always being at the end of the phone. Thanks to Paul for his patience with the final push.

I would like to give thanks to the MS Trust for funding the clinical trial and to The School of Health Professions funding my PhD. Above all I would like to give thanks to all the people who volunteered their time and efforts to participate in the trial and the therapists who gave up their time for the trial. Without these people this research would not have been possible.

AUTHOR'S DECLARATION

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Graduate Committee.

Work submitted for this research degree at the Plymouth University has not formed part of any other degree either at Plymouth University or at another establishment

This clinical trial was financed by a charitable grant from the MS Trust and carried out in collaboration with the School of Health Professions Faculty of Health and Human Sciences, Plymouth University.

Relevant scientific seminars and conferences were regularly attended at which work was often presented and papers prepared for publication.

Publications to date

- Fox, E., Hough, A., Gear, M., Creanor, S. and Freeman, J. (2014) Pilates based core stability training in ambulant individuals with multiple sclerosis: a multicentre, blinded, randomised, placebo controlled trial. *Way Ahead,*18 (1) 6-8

- Freeman, J., Fox, E., Gear, M., Hough, A. (2012) “Pilates based core stability training in ambulant individuals with multiple sclerosis: protocol for a multi-centre randomised controlled trial.” *BMC Neurology*, 12(1)19

.

Presentations and Conferences attended

Posters

- Ultrasound imaging of the deep abdominal muscles in people with multiple sclerosis: a comparison with matched controls. Rehabilitation in Multiple Sclerosis (RIMS) conference, Brighton June 2014

- The effects of Pilates upon deep abdominal muscle activity in people with multiple sclerosis: an exploratory ultrasound study. Rehabilitation in Multiple Sclerosis (RIMS) conference, Brighton June 2014

- Pilates based core stability training in ambulant individuals with multiple sclerosis: a multicentre, blinded, randomised, placebo controlled trial. Plymouth Hospitals NHS Trust conference, Plymouth (first prize won) September 2013

Platform Presentations

- Ultrasound of the deep abdominal muscles. RIMS conference, Brighton, UK June 2014

- Core stability training as rehabilitation for multiple sclerosis Consortium of Multiple Sclerosis Centers (CMSC) conference Dallas Texas, USA, May 2014 (Invited Speaker)

- Ultrasound imaging of the deep abdominal muscles. Plymouth University Post Graduate Conference. UK, November 2013

- Pilates based core stability training in ambulant individuals with multiple sclerosis. European Conference for Treatment and Research in Multiple Sclerosis. (ECTRIMS) Copenhagen October 2013.

- Core stability and multiple sclerosis. MS Trust Advanced Study Skills Day. Leeds UK. September 2012 (Invited Speaker)

Word count of main body of thesis: 69,475 (22nd February 2015)

Signed…….

Date……..

List of Abbreviations (with associated explanations)

10mtw: Ten metre timed walk

ABC scale: Activities Balance Confidence Scale

Abdominal drawing in manoeuvre (ADIM): Voluntary contraction of the deep abdominal muscles in order to stabilise the spine, achieved by drawing the navel towards the spine (McGalliard *et al* 2010).

ACSM: American College of Sports Medicine

Activation: Muscle activity, synonymous with recruitment when used to describe muscle activity.

Active straight leg raise (ASLR): a biomechanical test to asses load transfer between the legs and spine via the pelvis. During the ASLR a subject lies supine with knees in extension and lifts the leg to between 5-20cm (Gatti *et al* 2008; Teyhen *et al* 2009).

ADL: Activities of Daily Living

BBS: Berg Balance Scale

BF: Biceps Femoris

CI: Confidence intervals

COP: Centre of pressure

CNS: Central nervous system

CT: Computerised Tomography

EDSS: Extended Disability Status Scale

EMG: Electromyography

EO: External Oblique

ES: Erector Spinae

FR: Functional reach

FRT: Functional Reach Test

FFRT: Forward Functional Reach Test

ICC: Intraclass correlation coefficient

IO: Internal Oblique

IAP: Intra-abdominal pressure

LBP: low back pain

LOCF: Last observation carried forward

LFRT: Lateral Functional Reach Test

MCID: Minimum clinically important difference

MDC: Minimal detectable change

MRI: Magnetic resonance imaging

MS: Multiple sclerosis

MSWS-12: 12 Item Multiple Sclerosis Walking Scale

MVC: Maximal voluntary contraction

Pilates: A programme of exercises developed by Joseph Pilates in early 1900's, with an emphasis placed on trunk stabilisation, breath control and precision of movement. Now commonly adopted for clinical use (Wells *et al* 2012)**.**

US: Ultrasound

USI: Ultrasound imaging

RA: Rectus abdominis

RCT: Randomised controlled trial

RF: Rectus femoris

Recruitment: Synonymous with activation (in the context of abdominal muscle activation)

SIJ: Sacro-illiac joint

SD: Standard deviation

SEM: Standard error of measurement

‘**The Clinical Trial’:** The multicentre randomised controlled trial performed to assess the effects of Pilates based core stability exercises upon the balance and mobility of people with MS. This is the main study of this doctoral thesis and is referred to throughout as the clinical trial.

TrAb: Transversus Abdominis

TUG : Timed Up and Go

T25FWT: Timed 25 foot walk test

VAS: Visual Analogue Scale

Chapter Overview

Section One: The clinical trial

The first section introduces the concept of core stability, provides a literature review of the effects of Pilates and the outcome measures used. It details the methods and results of the clinical trial and discusses the results. The chapters within this section are detailed below.

Chapter One: Provides an overview of the clinical course of MS, balance and mobility impairment as consequence of MS according to the ICF, the effects of exercise upon people with MS and the rationale for performing the trial.

Chapter Two: Is a literature review of the concept of core stability. It defines core stability, identifies problems with assessing core stability and discusses the contribution of the deep abdominal muscles to balance.

Chapter Three: Is a literature review of the effects of Pilates and core stability training exercises upon balance and mobility, taking into account both healthy and clinical populations.

Chapter Four: Provides a rationale based on the literature, for the choice of outcome measures used in the clinical trial.

Chapter Five: Describes the methods used in the clinical trial.

Chapter Six: Reports the results of the clinical trial.

Chapter Seven: Provides a discussion and explanation of the clinical trial results with a summary.

Section Two: Ultrasound imaging (USI) of the deep abdominal muscles

The second section reports and discusses all the literature review and research findings of the exploratory USI study. The chapters within this section detailed below.

Chapter One: Introduces the use of USI as a method of measuring the deep abdominal muscles and provides a literature review of its psychometric properties.

Chapter Two: Reports on the reliability study performed prior to using USI in the clinical trial.

Chapter Three: Reports the findings of the study 'USI of the deep abdominal muscles of people with MS: a comparison with matched controls', which was performed alongside the clinical trial.

Chapter Four: Reports the findings of the study 'the effects of Pilates upon the deep abdominal muscles of people with MS' which was performed as part of the clinical trial.

Chapter Five: Reports the findings of correlations between Functional Reach scores and USI of the deep abdominal muscles.

Chapter Six: Summarises the literature and research finding of USI of the deep abdominal muscles

Chapter Seven: Provides directions for future research

Section One: The Clinical Trial

Chapter One: Introduction to the Clinical Trial

1.1 Introduction to Pilates based core stability training for people with MS

Pilates is a form of exercise which has grown in popularity over the last two decades (Wells *et al* 2012). The system of exercises was designed by the late 'Joseph Pilates' during the First World War and was influenced by gymnastics, yoga and tai- chi (Siler 2000). The Pilates system of exercises places a heavy focus on training the deep abdominal muscles in order to attain 'core stability'. The intention being that a stronger more stable core will result in improved outcomes in terms of balance and mobility (Bird *et al* 2012; Bird and Fell 2013) and pain reduction (Wajswelner et al 2012). Pilates originally gained popularity within the dance community, in more recent years, the exercises have been adapted and modified to be used in clinical populations. Training courses have been established in order to train therapists to apply the concepts of Pilates in clinical practice (Tulloch *et al* 2012).

In neurological rehabilitation the concept of achieving trunk control to assist in balance and mobility has been a central tenant of the Bobath approach (Smedal *et al* 2006) and has been used by therapists for the last 40 years (Raine *et al* 2009). Pilates based core stability training exercises have more recently been used by therapists working with people with MS in order to improve outcomes (Freeman *et al* 2010). In addition people with MS have been reported to enjoy this form of exercise (van der Linden *et al* 2013) and self-finance attendance at Pilates classes. To date there have been four studies performed using Pilates interventions in MS, three of which were published after designing the trial.

However there is no published conclusive evidence to evaluate the effectiveness of Pilates for improving balance and mobility in this population.

1.2 Multiple sclerosis: epidemiology and pathophysiology

This section will briefly define the epidemiology, aetiology and clinical course of MS and then consider the impact of MS upon balance and mobility according to the International Classification of Functioning, Disability and Health Framework (ICF).

Multiple sclerosis is considered to be the most common degenerative neurological condition affecting young adults with a prevalence of 110 cases per 100,000 in the population (Mackenzie *et al* 2014) , and an incidence of 1.12 to 6.96 per 100,000 in the European population (Alcalde-Cabero *et al* 2013).

The exact aetiology remains unclear (Asano *et al* 2009) but it is considered to be resultant from genetic susceptibility of an individual, combined with an environmental trigger (Compston and Coles 2008). Combined, these produce a succession of events resulting in acute inflammatory injury of the nerve, axons and glia, resulting in neuro- degeneration (Lassmann *et al* 2012). Sclerotic lesions can occur in any myelinated structure in the central nervous system (CNS), with a predilection for white matter tracts (DeLuca *et al* 2004). Involvement of the motor, cerebellar, sensory, visual tracts and vestibular apparatus and cognitive structures can occur (Freedman *et al* 2013). Deficits in these areas can result in motor, sensory and proprioceptive impairments, many of which can occur in a single person to varying degrees. These physiological impairments have consequences for people with MS and can result in problems

with both balance and mobility even in the early stages of disease onset (Martin *et al* 2006; Pike *et al* 2012).

Clinical course and subtypes of MS

The clinical course of MS varies between individuals. There is a pattern characterised by acute periods of exacerbation (relapses) which can lead to a gradual deterioration in neurological function (Lublin and Reingold 1996; Polman *et al* 2011). It is now considered that there are four main subtypes of MS; relapsing- remitting MS, primary progressive MS, secondary progressive MS and progressive relapsing MS (Lublin *et al* 2014) however there is often a lack of clarity in distinctly defining the subtypes. Benign MS and malignant MS and clinically isolated syndrome have been described as further subtypes (Lublin and Reingold 1996; Lublin *et al* 2014). The subtypes are briefly described below, as originally defined by Lublin & Reingold (1996) and then later revised by Lublin *et al* (2014); a detailed description and discussion of these phenotypes is beyond the scope of this thesis.

Relapsing- remitting MS: disease course has clearly defined relapses with either full recovery or leaving some residual neurological deficit.

Primary progressive MS: disease progression from onset with occasional plateaus but no district relapses.

Secondary progressive MS: initially a relapsing remitting course followed by progression with or without occasional relapses

Progressive relapsing MS: no consensus definition however characterised by a combination of relapses and progression.

Treatment of MS with disease-modifying drugs aims to reduce both the frequency and severity of attacks and to lessen disease progression (Freedman *et al* 2013). However, despite many pharmacological interventions being available there is no known cure for MS (Lassmann 2011). The socio-economic impact of walking and balance impairments is significant for people with MS (Pike *et al* 2012); physical therapy interventions are used to address these issues (Motl *et al* 2010; Paltamaa *et al* 2012).

In summary MS is characterised as an auto-immune degenerative neurological condition. It is one of the most prevalent neurological diseases affecting young adults. The aetiology remains unclear and the clinical course is varied.

1.3 The International classification of Functioning, Disability and Health (ICF)

The World Health Organisation (WHO) published the ICF as a conceptual framework for the definition and measurement of health and disability (WHO, 2001). The WHO not only recognises the burden of long term health conditions, but also the importance of focus upon function. The ICF categories can be used as a starting point for objectification of well-being. Additional benefits of using this conceptual framework are that it has worldwide cultural applicability and is integrative, neither medical nor social. The ICF can be used to assist in clinical research and intervention studies by optimising the comparability of results (Cieza and Stucki 2008). The ICF has been criticised however for being difficult to make clear distinctions between activities and participations when considering mobility (Paltamaa *et al* 2008). In this thesis impairments in mobility and balance and the interventions used have been considered in light of the ICF

conceptual framework. This section will address the impact MS has upon both mobility and balance.

1.3.1 The effects of MS upon walking

ICF definition of walking

According to the ICF definition 'walking' (code d450) is defined as 'moving along a surface on foot, step by step, so that one foot is always on the ground, such as when strolling sauntering, walking forwards, backwards or sideways and includes walking short or long distances, walking on different surfaces, and walking around obstacles, but excludes transferring and moving around' (Cieza & Stucki 2008 page 307). In this thesis the term walking will be used to describe walking as defined above, and considered an aspect of mobility.

Walking impairments

An estimated 75% of people with MS report problems with walking (Swingler and Compston 1992) and surveys indicate that this is a major concern for people with MS (Heesen *et al* 2008). Many physiological factors can influence walking, including; motor impairments such as lower limb and trunk neuromuscular weakness (Yahia *et al* 2011); cerebellar ataxia (Cameron *et al* 2008); fatigue (Smith *et al* 2011); sensory impairments such as visual symptoms and reduced sensation (Van Emmerik *et al* 2010). These in addition to psychosocial issues regarding anxiety and loss of confidence (Newsome *et al* 2011), either in isolation or in combination can result in problems walking.

Walking difficulties can lead to a cycle of inactivity. The subsequent deconditioning associated with this is typically accompanied by a further decline in ability (Dalgas and Stenager 2012). Walking is therefore considered to be one of the most important goals in neurological rehabilitation (Holland *et al* 2006) and a paramount aim of physiotherapy for people with MS (Paltamaa *et al* 2008).

1.3.2 The effects of MS upon balance

ICF definition of balance

The ICF definition of balance encompasses 'changing and maintaining body position' (ICF code d410- d 429), and is categorised in the mobility domain of the activities and participation component. For the purposes of this thesis balance can be described as 'maintaining a standing posture' (ICF code d4154) as this reflects the limits of stability whilst standing (WHO, 2001). In order to maintain an upright posture the integration of multiple sensorimotor processes are required (Prosperini *et al* 2011). The ability to generate co-ordinated movements and maintain the centre of mass within the limits of stability are crucial to maintaining balance (Shumway-Cook and Wollacott 2001).

The maintenance of balance is essential to function, and is an integral component of many Activities of Daily Living (ADL's) (Paltamaa *et al* 2007). Balance is closely related to the nature of the task to be undertaken and the environment in which it is performed (Paltamaa *et al* 2012), and the interaction with environment is accounted for in the conceptual framework of the ICF.

Balance impairments

Balance impairments present a significant problem for people with MS (Frzovic *et al* 2000) and may contribute to advancing disability (Hebert and Corboy 2013). Impairments in the vestibular (Hebert *et al* 2011) visual (Kesselring 2010), motor (Newsome *et al* 2011) and somatosensory systems (Cameron *et al* 2008) occur as a consequence of the central nervous system (CNS) damage which occurs in MS. Pathological lesions detected in the brainstem and cerebellum can interfere with sensory integration and contribute to impaired postural control (Prosperini *et al* 2011); as may spasticity (Sosnoff *et al* 2010). Any one of these factors in isolation, or in combination, can significantly impact upon balance. Even in the early stages of the disease, impaired balance has been demonstrated in people with MS in comparison to age and gender matched healthy controls, even in the absence of clinical disability as determined by routine clinical assessment (Martin *et al* 2006).

Impaired balance has consequences for people with MS, and has been reported to correlate with increasing disability (Boes *et al* 2012), memory and cognitive impairments (D'Orio *et al* 2012) and reduced mobility (Frzovic *et al* 2000). The incidence of falls has been found to be significantly higher than in matched controls and the incidence of injurious falls is greater still (Coote *et al* 2013). In considering that fear of falling has been found to curtail activity (Gunn *et al* 2013) and the higher incidence of injurious falls in those with impaired balance, management of this should become a clinical priority.

In summary, MS affects balance and mobility even in the early onset of disease. Impairments in balance and mobility have consequences and are associated with advancing disability and socio-economic impacts.

1.4.1 Exercise for people with MS

Scientific evidence supporting the beneficial effects of exercise is indisputable and outweighs the potential risks in most adults (Garber *et al* 2011). Studies which compare levels of physical activity between people with MS and other chronic diseases show that physical activity is particularly low in people with MS (Motl *et al* 2005). The incidence of osteoporosis (Nieves *et al* 1994), depression and death from cardiovascular disease is increased in the presence of MS (Brønnum-Hansen *et al* 2004), which is thought to be associated with inactivity and lack of ability to perform physical functions (Dalgas *et al* 2008). Furthermore inactivity is associated with atrophy and loss of muscle strength which can have negative implications upon functional capacity (Dalgas *et al* 2008) and quality of life (Marck *et al* 2014).This section will describe the effects of exercise for people with MS.

In previous years people with MS were advised not to participate in physical activity. This was in part because of symptom instability in response to increased core temperature (White *et al* 2000). Furthermore, it was proposed that avoiding exercise would preserve energy and decrease fatigue (Dalgas *et al* 2008). However, research suggests that the exacerbation in symptoms experienced by people with MS is temporary and normalised within 30 minutes of exercise cessation in 85% of people (Smith *et al* 2006). There is now a a growing body of scientific evidence to indicate that engaging in appropriate

structured exercise as part of rehabilitation is of benefit to people with MS, for improving function, quality of life and fatigue (Latimer-Cheung *et al* 2013). Exercise interventions have been found to have a positive effect upon body function, structure, activity and participation according to the ICF frame work (Rietberg et al 2004; Asano *et al* 2009). Not only has exercise been found to be beneficial, but the safety of exercise for MS has been established (Rietberg *et al* 2004; Dalgas *et al* 2008; Pilutti *et al* 2014) and may even reduce rate of relapse (Marck *et al* 2014; Pilutti *et al* 2014). In light of the fact that no pharmacological intervention has proven to effectively modify long term disease progression in people with MS (Mantia *et al* 2013; Cross and Naismith 2014), continued research to determine the effects of different types of exercise therapy as safe and efficacious methods of modifying progression is justified. As a consequence, many people with MS seek information as to the type, frequency, duration and intensity in which to perform exercise to gain maximum benefit (Asano *et al* 2009).

Physiotherapy exercise interventions have been used in order to address impairments in balance (Paltamaa *et al* 2012) and mobility (Snook and Motl 2009) for people with MS. Exercise has been found to result in a small yet clinically significant improvement in balance (Paltamaa *et al* 2012) and improvements in mobility (Motl *et al* 2010; Latimer-Cheung *et al* 2013) however the most beneficial dose of exercise is yet to be established (Collett *et al* 2011). Uncertainty has existed for many years regarding the most appropriate type of exercise for people with MS (Karpatkin 2005), with a paucity of research which assesses popular exercise (such as yoga and Pilates) highlighting the need for high quality RCTs to be performed (Latimer-Cheung *et al* 2013).

1.4.2 Type of exercise

In terms of type of exercise, resistance and endurance training have been most extensively investigated in people with MS. More specifically, resistance training (exercising muscles against resistance, using weights, bands or body weight), has been deemed as a safe, well tolerated and effective method of improving strength for people with mild to moderate MS (Dalgas *et al* 2008; Latimer-Cheung *et al* 2013). Similarly endurance training has been found to be a safe method of improving walking distance (Dettmers *et al* 2009). For people with more advanced disease and/or with symptoms such as marked spasticity, ataxia, weakness and fatigue, the ability to engage in traditional resistance programmes may simply not be possible (Karpatkin 2005). Pilates is composed of a series of exercises, which could be classified as low intensity resistance exercise (by using body weight as resistance), in addition to balance and coordination exercise (McNeill 2014).

1.4.3 Deconditioning and reversibility of impairments

The physical impairments noted in people with MS may be the result of either disease progression (i.e. demyelination and axonal degeneration) (Cameron and Wagner 2011) or as a result of secondary deconditioning from reduced physical activity (Motl *et al* 2010). The degree to which impairments are reversible is uncertain and until recently it was considered that muscle atrophy, loss of stamina and endurance as a result of reduced physical activity could be addressed whereas impairments from underlying neuronal degeneration were permanent (Dalgas *et al* 2008). However research has suggested that exercise

may have an anti-inflammatory disease modifying effect (Le Page *et al* 1996; White & Castellano 2008; Golzari *et al* 2010).

Summary of the effects of exercise in MS

Despite an exponential increase in the number of studies evaluating exercise in recent years, uncertainty continues to exist within the published evidence base regarding the optimal type and dose of exercise required to generate improvements in balance and mobility in people with MS. Additionally researchers performing systematic reviews, meta-analyses and guideline developments report that there is a paucity in high quality research to inform practice about some commonly used exercise interventions, such as Pilates (Karpatkin 2005;Dalgas *et al* 2008; Asano et al 2009;Latimer-Cheung *et al* 2013). Hence there is a requirement for high quality, adequately powered randomised controlled trials to be performed to address this.

1.5 Pilot research and design of the clinical trial

The concept of this clinical trial was based upon a national call for research questions to the 'Therapists in MS' (TiMS) group in 2008, in order to address questions raised by therapists working in clinical practice. In response to this, pilot research was designed to investigate the effect of Pilates based core stability exercises (heron referred to as Pilates) upon the balance and mobility of people with MS. The pilot research was performed as a multi-centre pragmatic series of single case studies to explore the feasibility and preliminary effectiveness of Pilates, and responsiveness of the outcome measures used

(Freeman *et al* 2010). Based on the results of this pilot study, a powered, assessor blinded multi-centre randomised controlled trial (RCT) was designed and implemented to determine the effectiveness of Pilates for improving balance and mobility in people with MS.

Given the time and resources required to perform the trial the opportunity was taken to evaluate the effects of a programme of standardised physiotherapy exercises (heron referred to as Standard Exercise {SE}), based on those used by Barrett *et al* (2009) and considered at that time to be reflective of NHS clinical practice. Finally, one of the recognised aims of Pilates is to selectively target the deep abdominal muscles in order to optimise the stabilising effect (Queiroz *et al* 2010). To investigate the effect of Pilates at the level of impairment, an exploratory ultrasound imaging (USI) study of the deep abdominal muscles was performed on a subset of participants. The aims of the trial were published in a peer reviewed protocol (Freeman *et al* 2012).

1.6 Aims of this clinical trial

Building upon the published pilot study, the primary aim of this clinical trial was to compare the effectiveness of a 12 week programme of individualised face to face Pilates sessions with a Relaxation exercise (placebo- control).

Secondary aims were to: 1) compare a 12 week programme of Standard Exercise, with the Relaxation- placebo, 2) compare the Pilates programme with the Standard Exercise programme, and 3) use USI to explore if changes in

resting thickness and activation levels of these deep muscles occur following exercise intervention.

Section One, Chapter Two: The concept of core stability

2.1 Introduction

The ability to stabilise the trunk in order for independent limb movement to occur has been a cornerstone of neurological therapy since the evolution of the Bobath concept, originating nearly 40 years ago (Raine *et al* 2009). Therapy based on the Bobath concept has been used to improve balance and mobility of people with MS (Smedal *et al* 2006). In more recent years core stability exercise programmes have been used, as part of physiotherapy interventions, to improve balance and mobility in MS (Freeman *et al* 2010).

The concept of core stability was proposed by Panjabi (1992) and since then there has been a plethora of research performed regarding the role of the deep abdominal muscles in spinal stabilisation. Contributions of the deep abdominal muscles to trunk stability in the neurologically impaired person has not been widely researched.

The purpose of this chapter is to explain the theoretical underpinning behind the use of core stabilisation programmes and how neuromuscular spinal stabilisation is achieved. In addition problems associated with the classification and measurement of core stability will be discussed.

2.2 Core Stability

2.2.1 The concept of core stability

In the early 1990's Panjabi proposed a theoretical model of 'core stability' based on the interdependence of three subsystems: the non- contractile tissues (osseous and ligamentous spine), the contractile (muscle) and the neural

system (Panjabi 1992) (see figure 1). This proposed theoretical model has influenced two decades of research and clinical practice (Hoffman and Gabel 2013). Despite this, there remains a lack of conclusive evidence demonstrating significant clinical benefits resultant from core stability training (Wajswelner *et al* 2012b). Many misinterpretations and misconceptions have arisen out of the published research and consequently the definition of 'core stability' remains without universally accepted consensus (Borghuis 2008). Definitions will be examined in section 2.2, page 43.

Much of the research has focused on the stabilising role of TrAb and has possibly over-emphasised the importance of training this muscle alone (Lederman 2010). The term core stability has almost become synonymous with TrAb training (Brooks 2012) and up until recently it has been commonly considered that TrAb can be isolated (Hodges & Richardson 1999) and retrained (Herrington and Davies 2005). As neuromuscular control depends on complex synergy between anatomical structures and neural control, the ability to isolate TrAb activity is now considered unlikely (Lederman 2010; Morris et al 2013).

Core stability is dependent on the co-activation and co-ordination of trunk muscles and is reliant upon sensory motor control (Morris *et al* 2013). A recent and encompassing theoretical model was proposed by Hoffman and Gabel (2013) which integrates elements of stability and mobility. Figure 2 demonstrates how each of these subsystems interact with neural and non-neural elements, presenting a more sophisticated model of core stability than the original model proposed by Panjabi. Hoffman and Gabel (2013) suggest that the ability of both stabilising and mobilising systems to work in harmony will

subsequently determine quality of movement. They propose that stability and mobility systems are separate but act in an integrated way under neural control. This requires a synergistic relationship between dependent neuromuscular components. Conversely the malfunction of either system will negatively affect all other subsystems and consequently efficiency of movement (Hoffman and Gabel 2013). Malfunction within these systems may be a result of pathological lesions within any of the subsystems.

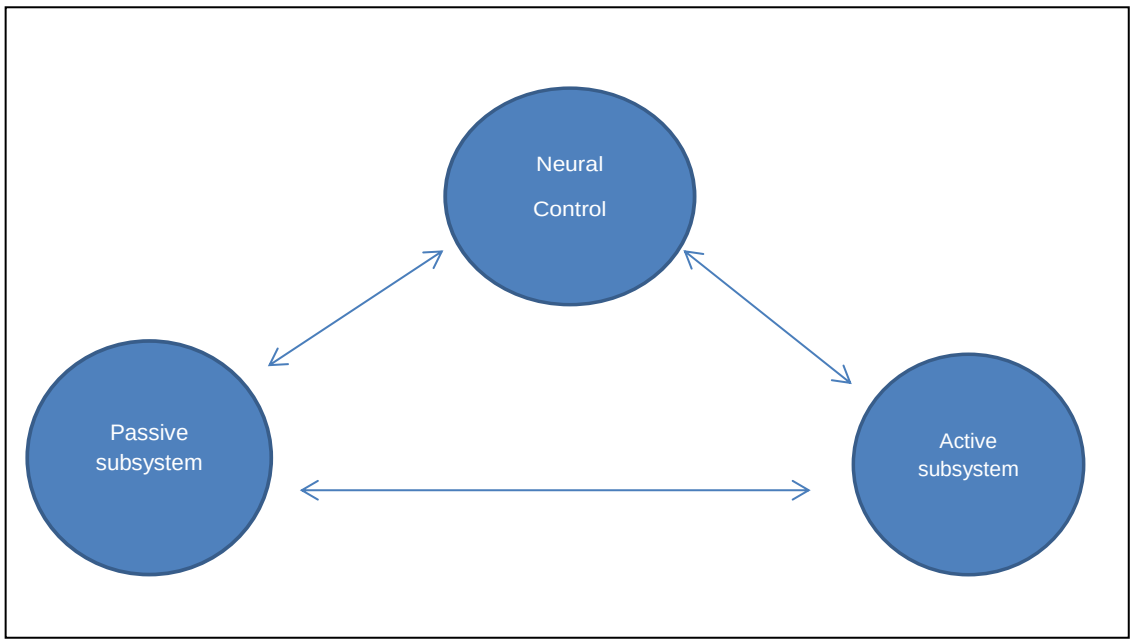


Figure 1: Panjabi's model of core stability

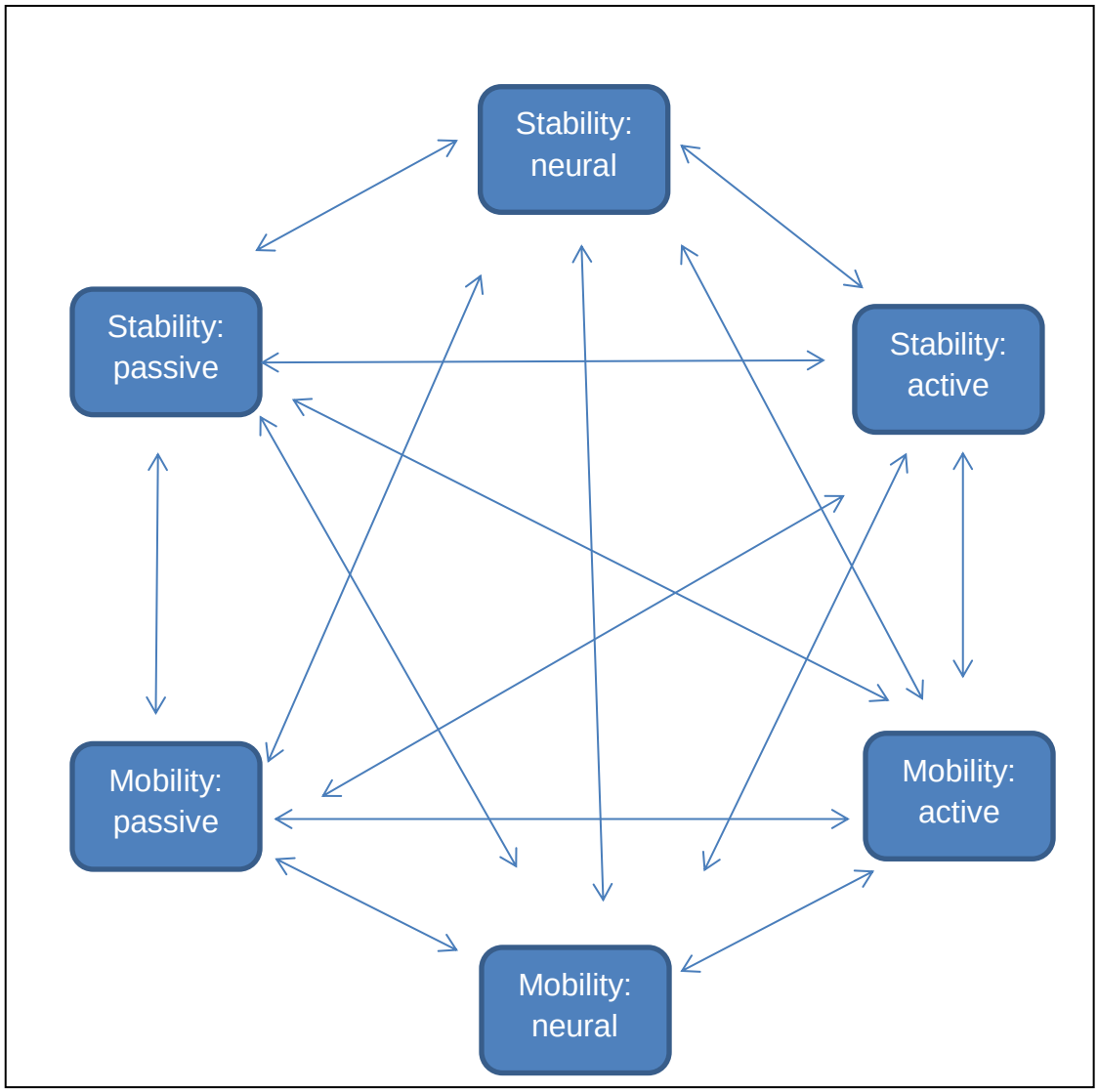


Figure 2: Hoffman and Gabel's model of core stability

which demonstrates interconnections between stability and mobilising subsystems (2013), adapted from Hoffman and Gabel (2013).

Concurrent to these theories, Key (2013) suggested that 'the core' contains three interdependent functions which include: breathing; postural control mechanisms of the axial column; and postural control in response to movement of the limbs. Iscoe (1998) implicated that TrAb additionally acts as a respiratory

muscle. Further to this Wallden (2013) proposed that any stressor, either in isolation or summation will create a sympathetic response resulting in an increased respiratory rate. He suggested that when breathing rate is accelerated, accessory musculature becomes preferentially recruited resulting in compromise of the diaphragm and TrAb via the autonomic nervous system. This theory has been further expanded by Key (2013) who suggests that breathing can transiently change the volume and shape of the trunk, creating slight postural disturbances which are corrected. Empirical evidence for the role of TrAb as an accessory muscle of respiration has been performed (DeTroyer et al 1990; Smith *et al* 2009), and while it is well established that TrAb is affected by respiration (Iscoe 1998), the link between respiration and the clinical application of core stability remains largely theoretical.

Whilst theories of core stability have some supporting empirical evidence, research has been mainly performed in healthy people or people with LBP, leaving many assumptions regarding the effect of core stability and core dysfunction on the balance and mobility of people with neurological conditions, such as MS.

The concept of trunk stability in the field of neurological rehabilitation is certainly not a recent proposal with alignment of the trunk and the ability to move limbs from a stable base being a central tenet to the Bobath concept, dating back to the 1960's (Raine *et al* 2009). In the absence of any rigorous studies evaluating core stability exercise, the concept of retraining the deep abdominal muscles in order to improve function in the neurologically impaired person is little more than anecdotal.

Whilst there have been studies which demonstrate that core stability training (including Pilates) can improve function in healthy people (fully explored in chapter three page 66) , there is no current evidence to determine the mechanisms of why a stronger and more stable trunk improves balance and mobility, or indeed if a stronger trunk is responsible for these gains (Granacher *et al* 2013). Theoretically, a stronger 'core' provides proximal stability in order that isolation of the limbs for distal mobility may be attained, fitting with the Bobath concept of physiotherapy. In support of this Ferreira *et al* (2010) found that difficulty with functional activities such as rising from a chair and stair negotiation was correlated with a poor ability to recruit TrAb in people with LBP. It is not unreasonable to purport that this may also be the case for people with MS.

In summary, there are proposals surrounding the mechanisms of core stabilisation, the majority of which are narrative opinion based reviews. In order to ascertain the contributions of the abdominal muscles to lumbar stability and the effect that this has upon the trunk and consequently balance, the empirical research published needs to be evaluated.

2.2.2 Defining of core stability

In determining what is meant by the term 'core stability' there are several definitions which encompass various anatomical structures. Pilates style exercise has become commonly associated with, and sometimes even synonymous with, core stability training. Hence it is worth highlighting that Pilates was not originally intended to be a clinical intervention for improving core stability (Wells *et al* 2012). In his original definition of what is now deemed as

‘the core’, Joseph Pilates referred to the trunk as the ‘power house’ of the body. This included the gluteal muscles, all the abdominal muscles and the paraspinal muscles (Siler 2000). A different definition by Chek suggests that 'if one were to pull off the extremities (limbs) the core would be left’ (Chek 1998, in Wallden 2013b page 240). This definition acknowledges the contribution of digestion and respiration. The definitions of the core which include a larger proportion of the proximal anatomy are comparative to trunk stability, which is more commonly measured and quantified in neurological rehabilitation (Verheyden *et al* 2006).

Many of the definitions of core stability do not encompass the gluteal stabilising musculature which is important in trunk stability, and connected to the deep abdominal muscles via the thoracolumbar fascia (Borghuis 2008). Whilst these muscles are undeniably important in balance and mobility they will not be described in this thesis (due to word constraints), as the interventions of the RCT and the exploratory ultrasound (US) research undertaken for this thesis focus on training and measurement of TrAb and IO. Detailed descriptions of the hip and shoulder musculature can be referred to in Drake et al (2005).

In other texts ‘the core’ has been described as including a functional unit comprising of TrAb, pelvic floor, the diaphragm and multifidus (Richardson *et al* 1999); again this definition includes the effect of respiration. Spinal or lumbar stabilisation is often described instead of core stability. Spinal stability has been defined as ‘sufficient spinal stiffness to minimise unnecessary movement between spinal segments’ (Morris *et al* 2013). This provides a clear context as to the importance of studying spinal stability in LBP but does not take into account the influence of the rest of the trunk anatomy upon balance. Kibler et al (2006) defined core stability as the ability to control the position of the trunk over

the pelvis, which takes into account the entire neuromuscular interactions required for stabilisation. Reasons for these discrepancies remain unclear but could potentially be due to the research performed surrounding TrAb dysfunction in LBP which focus more closely on lumbar segmental stabilisation, whereas studies assessing balance consider the importance of the whole trunk.

In summary there is no published universal definition of core stability, which can make comparisons of studies difficult. For the purpose of this thesis, based on the available literature a definition of core stability has been proposed by the researcher (EF): “The ability to activate the deep abdominal muscles and surrounding trunk musculature in order to stabilise the lumbar spine and control the position of the trunk above the pelvis’’

2.3.3 Quantification of core stability

Valid and reliable quantification of task performance (such as the ability to stabilise the lumbar spine and consequently the trunk) is required if it is to be used as an outcome measure (Amato and Portaccio 2007). Despite the widespread focus upon interventions aiming to improve core stability, the classification and quantification of core stability remains poorly defined with little consensus on the use of valid and reliable measures (Borghuis 2008).

Some clinicians subjectively assess core stability ‘by eye’, by visually analysing the person’s ability to stabilise in differing positions (Weir *et al* 2010). This visual assessment of core stability is open to wide subjective interpretation. It is perhaps therefore unsurprising that inter and intra-observer reliability is poor, with ICC’s ranging from 0.09 (CI=0.01-0.21)-0.55 (CI= 0.35-0.66) (Weir *et al*

2010). An alternative method of evaluating core stability is for therapists to palpate the abdominal muscles to determine activation. Costa *et al* (2006) performed a study comparing the reliability of palpation to pressure bio-feedback for assessing TrAb activation in healthy young adults (n=29). They described both tests as achieving 'moderate intra-tester reliability' (palpation ICC: 0.52, CI 95% 0.29-0.75 and pressure biofeedback ICC: 0.58, CI 95% 0.28-0.78); inter-tester reliability was not determined. However, the ability to activate TrAb in isolation is questionable, as it is plausible that IO may be palpated and activated when using pressure biofeedback.

Other methods suggested for measuring core stability are also available, although none are specific in targeting the core stabilisers. Isometric dynamometry measures trunk muscle strength (Kibler *et al* 2006), however this does not give any indication of the onset of activation or the changes in spinal stiffness. Functional measures have been suggested such as timed single leg standing and single leg squats (Borghuis 2008), however their validity as measures of core stability could be questioned given the multiple interacting variables required to perform these tasks (e.g. lower limb strength, stability at multiple joints, sensory integration). It has been suggested that sitting balance may be a more appropriate measure of core stability (Cholewicki *et al* 2000; Preuss *et al* 2005) as this eliminates the effect of lower limb stability. Sitting balance has been applied for quantifying trunk stability in people with MS (Lanzetta *et al* 2004).

Kavcic *et al* (2004) assessed spinal stability using assessment of three dimensional lumbar motion, EMG of trunk muscles and calculated external forces to provide a precise biomechanical assessment of the effect of the

osseous, ligamentous and muscular structures in response to destabilisation. This however required expensive and invasive equipment and is therefore not feasible for use within routine clinical settings.

Borghuis (2008) suggested using the Sahrmann's scale of core stability (see table 1), in agreement with Akuthota and Nadler (2004) and used by Aggarwal et al (2010). The face validity of this measure however appears poor; ostensibly providing a progression of exercises to improve spinal stability rather than an assessment scale. Whilst one could argue that this scale provides a good clinical description of an individuals ability to stabilise and or position the trunk that could be helpful for assessment purposes, its psychometric properties have not yet been explored which significantly limits its usefulness for research purposes.

Level	Sahrmann's Lower Abdominal Exercise Progression
Base position	Supine with knees bent and feet on floor; spine stabilized with "navel to spine"
Level 0.3	Base position with 1 foot lifted
Level 0.4	Base position with 1 knee held to chest and other foot lifted
Level 0.5	Base position with 1 knee held lightly to chest and other foot lifted
Level 1a	Knee to chest (90° of hip flexion) held actively and other foot lifted
Level 1b	Knee to chest (at 90° of hip flexion) held actively and other foot lifted
Level 2	Knee to chest (at 90° of hip flexion) held actively and other foot lifted and slide on ground
Level 3	Knee to chest (at 90° of hip flexion) held actively and other foot lifted and slide not on ground
Level 4	Bilateral heel slides
Level 5	Bilateral leg lifts to 90°

Table 1 Sahrmann's scale of core stability / abdominal exercise progression

Adapted from Sahrmann (2002) in Akuthota and Nadler (2004)

In summary there is weak evidence to determine the reliability and validity of measures to assess core stability. It has been proposed that measurement of deep abdominal muscle activation, such as undertaken by USI, should be used when accuracy is required (Costa *et al* 2009). A thorough appraisal of the application, including the validity and reliability of using USI for this purpose is discussed in section two, page 232. Studies evaluating core stability training may additionally benefit from using reliable functional outcome measures to assess the effects of intervention in combination with impairment based measures assessing 'core stability'. This will enable the effects of the

intervention to be more comprehensively assessed and in a way that is meaningful to the patient.

2.3.1 Functional anatomy and relationship with trunk stability

A thorough knowledge of the anatomy of the abdominal musculature is helpful in understanding the biomechanics regarding activation. The abdominal wall covers a large area, spanning between the xiphoid process and costal margins superiorly and the iliac crest and public symphysis inferiorly (Drake *et al* 2005) (see figure 3).

The TrAb muscle originates from the iliac crest, inguinal ligament, thoracolumbar fascia, and costal cartilages (7-12), and inserts upon the xiphoid process, linea alba, pubic crest and pubis via conjoint tendon. TrAb is innervated by the thoraco-abdominal nerve (T6-T11), the subcostal nerve, (T12), the iliohypogastric (L1), and ilioinguinal (L1) nerves. IO originates from the inguinal ligament, iliac crest and the lumbodorsal fascia and inserts to linea alba, pubis (via conjoint tendon) and ribs 10-12. It is innervated by the thoracoabdominal nerve (T6-T11), the subcostal nerve (T12), and the iliohypogastric (L1) and ilioinguinal nerves (L1) (Drake *et al* 2005; Ger 2009; Willard *et al* 2012).

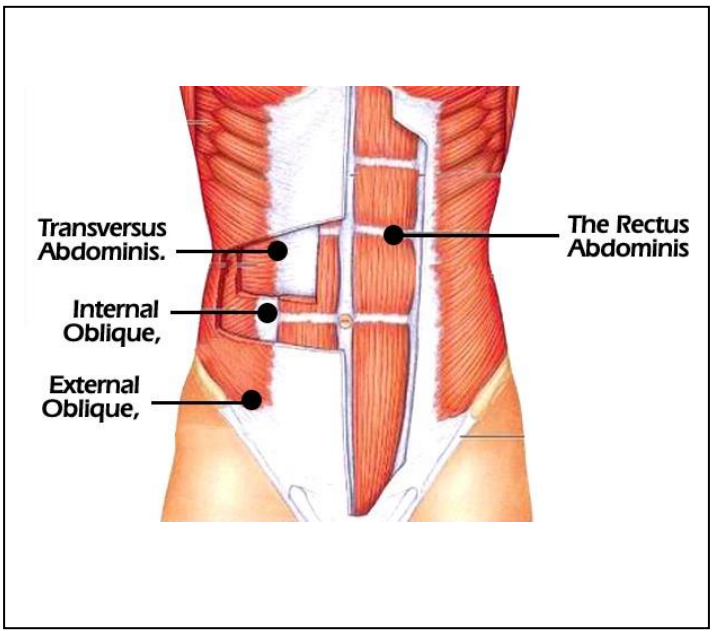


Figure 3: Abdominal muscles

(adapted from http://vancouverspinecarecentre.com)

The precise functional role of the deep abdominal muscles has been subject to considerable discussion; and research spanning almost 25 years has focused on the function of TrAb (DeTroyer *et al* 1990; Hodges 1999; Morris *et al* 2013). Primarily TrAb is thought to contribute to spinal stability by increasing intra-abdominal pressure (Beales *et al* 2009), and IO and EO are considered to rotate and flex the trunk (Drake *et al* 2005). Further evidence suggests that there is also a synergistic neuromuscular coordination and co-contraction of agonist and antagonist paraspinal, deep abdominal and trunk muscles, which facilitate spinal and furthermore trunk stabilisation in addition to the primary agonist role (Kavcic *et al* 2004; Morris *et al* 2013).

All of the deep abdominal muscles are considered to play a role in stabilising the spine, however, the relative contributions to this function remain unclear

(Gibbons & Comerford 2001; Urquhart *et al* 2005; Ainscough-Potts *et al* 2006). The following text evaluates the contributions of the deep abdominal muscles to spinal stability.

2.3.2 Activity in the core musculature

It is suggested that the core muscles perform different functions depending on location and muscle fibre type (Crisco and Panjabi 1991). A narrative review by Gibbons and Comerford (2001) classified the abdominal muscles as either local stabilisers, global stabilisers or global mobilisers, depending on their predominant function (see table 2). Categorising these muscles by function in this manner helps to gain an understanding of the anatomy of core stability.

Category	**Muscles**	**Action**
Local stabiliser	Transversus Abdominis and Multifidus	Stabilise the lumbar spine, increase intra-abdominal pressure
Global stabiliser	Internal and External Oblique	Flex and rotate the trunk, contribute to stability
Global mobilisers	Rectus abdominis and Erector Spinae	Flex and extend the trunk respectively

Table 2: Classification of the trunk muscles by Gibbons and Comerford (2001)

Gibbons and Comerford (2001) report that TrAb and multifidus do not contract to produce significant length changes within the muscle. Without changes in length muscles are unable to generate sufficient torque to act as agonists. This is evidenced by research which has demonstrated that TrAb does not produce length changes of more than 20% during lumbar flexion, extension and

rotational movements (McGill 1991) meaning that TrAb is not a prime trunk flexor but can contribute to trunk movement. In addition, EMG activity suggests that TrAb activity is continuous throughout movement (Hodges and Richardson, 1996). There is consensus opinion that TrAb is considered to be primarily a postural muscle (Gibbons and Comerford 2001) due to the composition of a higher percentage of tonic (slow twitch) muscle fibres, which are able to activate at low Maximum Voluntary Contraction (MVC) over long periods of time. This allows stabilisation of the lumbar spine and trunk (Haggmark and Thorstensson 1979).

Muscle fascicles originating from different structures have different functions and can assist in producing different movement. For example upper fibres of TrAb act to stabilise the rib cage, the middle fascicles stabilise the lumbar spine and the lower fibres support the abdominal contents and compress the sacroiliac joint (Urquhart *et al* 2005b). The implications for this study are that when performing USI, the placement of transducer over the muscle will provide a limited perspective on the activity of the muscle.

Multifidus is a deep intervertebral paraspinal muscle which contributes to spinal stability by acting as a tonic muscle. This is due to the higher percentage of type 1 (slow twitch/ tonic) muscle fibres than found in erector spinae (ES) and the fact that Multifidus is anatomically closer to the vertebrae (MacDonald *et al* 2006). Multifidus co-contracts with TrAb to increase spinal stiffness and spinal stabilisation and hence is important in assisting in core stability. Multifidus can also be reliably visualised with USI (Koppenhaver *et al* 2009), however the focus of this clinical trial was the role of the deep abdominal muscles in core stabilisation and hence multifidus was not measured in this dissertation.

3.3 The corset theory

Contraction of TrAb increases intra-abdominal pressure and stiffness of the spine by applying tension to the thoraco-lumbar fascia, thus increasing spinal stability. This phenomenon has consequently become known as the ‘corset theory’ of core stability (Richardson *et al* 1999), and has been applied clinically by therapists teaching patients to voluntarily activate the deep abdominal muscles with the use of an abdominal drawing in manoeuvre (ADIM) (Lim *et al* 2011).

Simulation of TrAb in cadavers has been shown to increase stiffness of segments of the lumbar spine, mainly by applying tension to the middle layer of thoracolumbar fascia (Barker *et al* 2006). This is most marked when the spine is in a neutral position and accounts for why therapists teach voluntary contraction of the deep abdominal muscles with the spine in a neutral position (Cruz-Ferreira *et al* 2013).

Activation of the abdominal muscles can be attained by either voluntary or automatic means. Voluntary activation is achieved by drawing in the navel towards the spine, termed the ‘abdominal drawing in manoeuvre’ (ADIM) (McGalliard *et al* 2010). Automatic activation is initiated by destabilising the spine with movement (see page 254).

2.3.4 Anticipatory feed forward activation of TrAb

Muscle strength alone does not explain the importance of the deep abdominal muscles in core stabilisation. In light of the fact that TrAb stabilises the spine

sufficiently for performing activities of daily living at MVC of 10% (Stokes *et al* 2011), sensory motor control is deemed an important aspect of core stability.

Core stability requires the use of sensory and motor processing strategies along with learned responses from previous experiences in order for anticipatory responses to occur. Stabilisation depends on three levels of motor control; spinal reflexes, postural responses modulated by the brain stem, and cognitive programming to produce appropriate muscle responses (Radebold *et al* 2001).

Spinal reflex pathways use proprioceptive input from muscles spindles and golgi tendon organs, the brainstem coordinates vestibular, visual and proprioceptive feedback in order to maintain postural control, and the cognitive programmes are based in stored central commands; which lead to voluntary adjustments (Shumway-Cook and Wollacott 2001). These anticipatory feed forward reactions allow the body to respond to perturbations created by mobilising (e.g. walking, arm movements) (Hodges & Richardson 1999) hence it becomes apparent that strengthening the core muscles alone may not be sufficient for retraining these muscles, giving rise to exercise programmes to improve motor control (Hodges 1999).

2.3.5 Levels of contraction

Core stability theories have been based on mathematical biomechanically engineered concepts of energy, stability and stiffness (McGill and Cholewicki 2001). Stiffening the spine increases the stability, however for efficient movement a dynamic equilibrium between stiffness and flexibility is required. With regard to this, only low levels of contraction of the trunk muscles are

required to give sufficient stability against minor perturbations (Borghuis 2008). Abdominal and paraspinal muscle contractions as low as 5% of maximum MVC are enough to provide stability of the spine for the performance of ADL, and 10% for vigorous activity (proposed by Kibler *et al* 2006 based on the theoretical modelling research of Cholewicki *et al* 1999), further supporting the importance of efficient motor control in providing core stability. Biomechanical modelling has demonstrated that whilst forced activation of 10% MVC of TrAb and IO increased spinal stability, increasing the forced contraction to 20% MVC did not further increase the stiffness of the lumbar spine. This supports the notion that only low levels of contraction are required to stabilise the lumbar spine (Stokes *et al* 2011).

2.3.6 Isolation of TrAb

Hodges and Richardson (1999) report that TrAb is activated in anticipation of movement, to provide stabilisation of the spine. Described as anticipatory feed forward reactions, Hodges and Richardson go on to suggest that there is a disassociation between the behaviour of TrAb and the other abdominal muscles, proposing that the motor command for TrAb activation may be independently controlled. A further study to support this demonstrated that when EMG is applied to differing regions of TrAb, the onset of TrAb activity in response to limb perturbation differs between regions. The lower fibres activated prior to middle and upper fibres with rapid arm flexion, however no difference was noted between the recruitment of TrAb middle and lower fibres and IO under these conditions (Urquhart, Hodges and Story 2005).

2.3.7 Discrepancies within the corset theory

Whilst empirical evidence exists demonstrating that destabilisation of the spine results in activation of the deep abdominal muscles (see above), Morris *et al* (2012) refute the 'corset hypothesis' with data which suggests that feed forward activation of TrAb is neither bilateral nor independent of the direction of arm perturbation. Their research suggests that TrAb activates in a diagonal pattern rather than a contralateral pattern. They propose that TrAb forms part of a synergy of muscle activity which contributes to axial rotational forces which act to oppose the direction of limb movement (perturbation). In rehabilitation, the ADIM is used as a method of voluntarily stabilising the spine (Herrington and Davies 2005), however Morris et al (2012) suggested that voluntary training of TrAb by use of the ADIM is not required to improve spinal stability (Morris *et al* 2013). This proposal is supported by a plethora of clinical research summarised by systematic reviews which fail to demonstrate that abdominal muscle training is superior to other exercise interventions for improving conditions such as LBP (Pereira et al 2011; Lim et al 2011) or impaired balance in MS (Marandi *et al* 2013).

Allison and Morris (2008), and more recently Morris et al (2013), proposed that there were methodological limitations in the research performed by the Hodges group in the 1990's (Hodges & Richardson 1999) to determine the role of TrAb, such as only using unilateral EMG and limited arm movements to create spinal perturbations.

2.3.8 The effect of posture upon spinal stabilisation

There is evidence to support the role of TrAb as a stabiliser of the lumbar spine in response to sudden external perturbations. This has been demonstrated using EMG in standing (Hodges & Richardson 1999), sitting (Urquhart, Hodges & Story 2005) and with gravity eliminated in side lying (Crommert and Thorstensson 2009). In side-lying the onset of TrAb activation was found to be independent to the direction of the trunk perturbation and either simultaneous or later than superficial abdominal muscles in lying (Crommert and Thorstensson 2009). This differs to the findings which report that TrAb activity is prior to the onset of movement in standing (Hodges & Richardson 1999). Urquhart, Hodges and Story (2005) reported that the recruitment of abdominal muscles differs depending on postural demands with recruitment of TrAb and IO delayed in sitting in comparison to standing. To summarise, posture, whether sitting, standing or lying may affect deep abdominal activation with evidence suggesting delayed onset in sitting.

2.3.9 TrAb and the role in respiration

The deep abdominal muscles activate during coughing, sneezing and vomiting by increasing intra-abdominal pressure (IAP) (Iscoe 1998). TrAb acts as an accessory muscle of respiration, and activation of TrAb has been found to increase IAP (Beales *et al* 2009) and expiratory effort (Kaneko *et al* 2006) . Original research performed by DeTroyer *et al* (1990) analysed the role of TrAb in respiratory function with EMG. Increased activity was demonstrated during forced expiration, coughing and laughing, however when breathing at tidal

volumes TrAb was not activated beyond the activity that was required to maintain sitting posture.

McGill & Karpowicz (2009) reported that heavy breathing whilst performing isometric core stability exercises, such as a counter poise in four point kneeling (see figure 4), did not increase activity in IO and EO beyond the MVC required to stabilise the spine during the exercise. They attributed this to the participants using the diaphragm, not the deep abdominal muscles during breathing. In conclusion, the deep abdominal muscles are important in contributing to respiration. The aim of this thesis is to evaluate the role of the core stability exercise programme upon the balance and mobility of people with MS, hence the effect of the abdominal muscles on respiration is only summarised here. The role of respiration upon the deep abdominal muscles and how USI is affected is discussed on page 241.

Figure 4: An example of counter poise in four point kneeling

2.3.10 Anomalies

Text book descriptions of anatomy may portray the deep abdominal muscles to be uniform amongst individuals. Cadaver studies demonstrate that this is not

the case. Urquhart *et al* (2005) performed a dissection study on 26 human cadavers to investigate the morphology of the deep abdominal muscles, namely TrAb, IO and EO. Each of these muscles has a number of primary osseous attachments. These include the costal cartilages, lumbar spine, iliac crest and pubis. Regional differences in orientation of TrAb and IO fascicles exist. Superior to the iliac crest, IO fascicles were orientated superior-medially, in contrast to fibres below the iliac crest which were horizontal. Five anatomical variations were identified; in TrAb there were cases of partial and complete detachment of TrAb from the iliac crest and an abrupt change of muscle orientation in the lower and middle regions with fusion of the lower fibres of IO (Urquhart *et al* 2005). Whilst one would not expect to find identical anatomy between individuals, this supports the findings of Kavcic *et al* (2004) and Morris *et al* (2013), who report considerable variance in the activation patterns of the deep abdominal muscles in order to stabilise the spine. This has implications for the reliable US measurement of these muscles

2.4. The trunk

2.4.1 Trunk stabilisation

The skeletal system is inherently unstable and requires the activation of anti-gravity muscles to generate constant tension to maintain a stable posture (Takayanagi *et al* 1995). The trunk is unstable without muscular control (Blaszczyk *et al* 1994) as the trunk responds to the influence of gravity (Lanzetta *et al* 2004). The CNS maintains the trunk position within spatial boundaries described as ‘stability limits’, which require both perception and the development of adequate postural responses to feedback from the visual,

vestibular and proprioceptive systems (Patton *et al* 1999). As the deep abdominal muscles comprise a significant proportion of the trunk musculature, it has been suggested that effective activation of the deep abdominal muscles influence activity of the entire trunk and consequently balance (Zedka *et al* 1998).

As previously discussed, the core muscles have been broadly categorised as local stabilisers, global stabilisers and global mobilisers (Gibbons and Comerford 2001). Whilst EMG data demonstrates that when ES and RA contract concentrically they produce large direction dependent movements of the trunk (such as extension and flexion respectively) (Kumar 2010), the notion that certain muscles act as stabilisers and others as mobilisers has been disproved by Kavcic *et al* (2004). In a study designed to provide a systematic biomechanical analysis to determine the role of the trunk muscles in response to destabilisation of the spine, Kavcic *et al* used highly sophisticated modelling to give detailed information about the role of the behaviour of trunk muscles during commonly prescribed stability exercises. Using EMG on 14 points of the trunk musculature, they determined the impact of artificial perturbation upon muscle contraction and spinal biomechanics in participants (n=10) performing exercises. Results of this study yielded some interesting findings, in that there was no consistent pattern across trunk muscles in their ability to affect stability of the spine. Contraction of quadratus lumborum, multifidus, and TrAb each created minimal changes to biomechanics of the lumbar spine, with IO and EO demonstrating the largest impact on spinal stability, irrespective of the task conditions.

In this experiment of human subjects no single muscle when activated at 0-100% MVC created destabilisation of the spine. Additionally no individual muscle reduced in activation during stabilisation. Also observed by Kavcic *et al* (2004) was the ability of muscles to change behaviour dependent on the exercise being performed. For example RA acted as an agonist prime mover during flexion based exercises, with ES acting as a stabiliser. These roles were reversed when lumbar extension exercises were employed. There is not, however, a consensus on this. For instance computer generated biomechanical modelling of the lumbar spine has suggested that forced activation of RA does not contribute to spinal stability (Stokes *et al* 2011).

In summary the research performed by Kavcic *et al* (2004) suggests that no single muscle is superior at stabilising the trunk. Consequently they recommend to train all of the trunk muscles if aiming to improve trunk stability. In support of this Morris *et al* (2013) reported that natural variance occurs in the muscle activation patterns which may reflect different strategies in stabilisation and mobilisation occurring in different people.

This research is counter to the proposal that in order to improve 'core stability', emphasis should be placed on the voluntary activation of TrAb (Hodges 1999) as taught in clinical Pilates (Owsley 2005), and that improving core stability depends on training the functional unit of the core stabilisers which have been described as TrAb, pelvic floor, multifidi and diaphragm (Richardson *et al* 1999).

4.2 Trunk muscle activation and balance

Trunk muscle activity has been studied in relation to balance in healthy people (Cetin et *al* 2008; Davidson *et al* 2009), and in those with in stroke (Karatas *et al* 2004) and MS (Lanzetta *et al* 2004). More specifically TrAb has been found to activate in synchronisation with erector spinae in response to load release perturbations (Crommert *et al* 2011). Activation in the other trunk muscles (IO, EO, RA and ES) as measured by EMG has been reported in response to sudden trunk perturbations in healthy people (Vera-Garcia *et al* 2007; *Jacobs et al* 2011) indicating that the deep abdominal muscles may contribute to balance.

Fatigue in trunk muscles has been associated with impaired balance in healthy young adults (n=30 mean age 24 years). Using an isokinetic dynamometer, trunk muscle strength (precise muscles not documented only 'flexor or extensor' muscles reported) was measured before and after exercise and correlated with dynamic balance. Trunk muscle fatigue (produced using isokinetic dynamometer) was weakly correlated with reduced dynamic balance (r=-0.37,p=0.45) (Cetin *et al* 2008). Fatigue in the lumbar extensor muscles has also been found to significantly impair balance recovery in response to perturbations as measured by centre of mass excursion (p=0.001) and centre of pressure trajectory (p=0.001) in healthy people (n=32) (Davidson *et al* 2009) further indicating that trunk musculature contributes to balance. A systematic review by Helbostad *et al* (2010) reported that fatigue in the trunk muscles induces postural instability during quiet standing and impairs functional reach tasks, further supporting the notion that the trunk muscles play an important role in balance. This potentially has important implications for people with other conditions where fatigue is an important symptom, such as MS.

2.4.3 Trunk muscle activity in people with neurological pathology

People with MS demonstrate delayed anticipatory postural adjustments, demonstrated as an impaired ability to activate the trunk and leg muscles prior to a forthcoming body perturbation (Krishnan *et al* 2012). EMG activity in RA and ES, in addition to biceps femoris, semitendinosus, soleus and tibialis anterior, was measured in response to repeatedly lifting a 2.27kg weight in people with MS (n=11 + 11 matched controls, EDSS > 5). People with MS displayed significantly delayed anticipatory muscle onset in ES ($p=0.01$) and a non-significant delay in RA ($p=0.09$). There was a reduced magnitude of anticipatory muscle activation in both RA and ES ($p<0.05$). All MS subjects demonstrated a smaller anticipatory centre of pressure in comparison to healthy controls ($p=0.001$). This study suggests that people with MS have reduced anticipatory muscle activity in the trunk muscles and delayed trunk muscle activation which the authors report may contribute to the reduction in stability, in terms of balance. These findings were noted even in people mildly affected with MS.

Evidence to support the importance of the trunk is also provided by other neurological conditions such as stroke. Dickstein et al (1999) performed a study to assess activity in the trunk muscles in people with stroke. EMG activity in the trunk muscles (RA and ES) was found to be reduced in hemi-paretic and hemiplegic patients post stroke (Dickstein *et al* 1999). Synchronous activation of these two muscles was greatest during voluntary dynamic tasks indicating their role as postural muscles in addition to acting as prime movers (Dickstein *et al* 1999). Furthermore, impairments in trunk muscle strength may affect balance in

people who have uni-hemispheric stroke (Karatas *et al* 2004). In a study which assessed trunk muscle strength and balance in stroke (n=38+ 40 matched controls) findings indicated that weakness in trunk extensor and flexor muscles, as measured by isokinetic dynamometry, was correlated with Berg Balance Scale scores (r= 0.32-0.64, $p<0.05$). The authors suggested that even mild weakening of the trunk muscles (undetectable by manual muscle testing) can interfere with balance and stability and increase functional disability. This study highlights the importance of trunk muscle strength upon balance in a neurologically impaired clinical population.

In summary, there is evidence to demonstrate that trunk muscles may affect balance in both the healthy population and people with neurological impairments, including MS.

2.5 Conclusion

Proposed theories of core stability involving the deep abdominal muscles originated in the early 1990's (Panjabi 1992) and have resulted in two decades of research surrounding the role of these muscles in spinal stability. Research has been performed which has focused on the role of TrAb in spinal stability and it has been proposed that the training of TrAb can improve core stabilisation. In previous years it was considered that delayed onset of activation was responsible for core dysfunction (Hodges and Richardson 1996). Recently questions about the role of TrAb acting as i) part of a corset of muscles to stabilise the spine and ii) activating prior to the onset of other muscles have been raised (Morris *et al* 2012). It is now acknowledged that all of the trunk muscles contribute to spinal stabilisation via a complex synergistic

neuromuscular coordination and co-contraction of agonist and antagonist paraspinal, deep abdominal and trunk muscles (Kavcic *et al* 2004). Further to this there appears to be no consensus in the literature of either the definition of core stability or reliable and valid methods of measuring core stability (Borghuis 2008). The majority of research regarding the role of deep abdominal muscles has been undertaken in people with LBP. There is a paucity of research in people with neurological conditions and this is particularly limited in regard to MS. Trunk stabilisation, rather than the measurement of specific muscles, is more commonly used as a measure of stability in neurologically impaired persons (Dickstein *et al* 1999). However when it is considered that the deep abdominal muscles comprise a significant component of the trunk anatomy, and it has been shown that impairment in these muscles may affect balance, research to determine the effects of the deep abdominal muscles upon both balance and mobility is justified.

Section one, Chapter Three: The effects of Pilates and core stability training upon balance and mobility: a review of the literature.

3.1 Introduction

In the 1990's research was performed suggesting that TrAb activation may be dysfunctional in the presence of low back pain (LBP) (Hodges and Richardson 1996). This, combined with proposed theories of core stability (Panjabi 1992) suggesting that voluntary activation of the deep abdominal muscles is required for lumbo-pelvic stability, resulted in a rise in the popularity of Pilates within the clinical rehabilitation setting. Pilates exercises were adapted and modified and courses were established in order to train physiotherapists to apply Pilates principles within clinical practice (Wells et al 2012).

Pilates uses a system of up to 50 simple repetitive exercises. All the Pilates exercises are based on the 'five essentials', which are described as breathing, cervical alignment, scapular and rib stabilisation, pelvic mobility and the use of the deep abdominal muscles (see table 3). Joseph Pilates believed that a strong trunk, was crucial to correct performance of the Pilates exercise repertoire (Muscolino and Cipriani 2004). Each exercise is initiated by voluntarily stabilising the core musculature including the abdominal, gluteal and paraspinal muscles and then proceeds through a controlled range of motion. Body weight is used as resistance, and changes in body position can be used to challenge participants (Kloubec 2011).

Traditional Principle	Definition
Centering	Tightening of muscular centre of the body or 'Powerhouse' by contracting the muscles located between the power house and rib cage
Concentration	Cognitive attention required to perform exercise
Control	Close management of posture and movement
Precision	Accuracy of exercise technique
Flow	Smooth transitions of movements within their sequence
Breathing	Moving air in and out of lungs in co-ordination with exercise

Table 3: The Principles of Pilates

(adapted from Wells *et al* 2012)

In neurological rehabilitation the concept of trunk stability is not modern, considered central to the Bobath concept which was first implemented in the 1960's (Raine *et al* 2009). Whilst anecdotal evidence suggests that neurological therapists often employ Pilates based core stability training as part of a management programme, the evidence base to support this is limited, with only four research articles published to date. It is however noteworthy that Pilates and core stability training are not synonymous. Pilates incorporates aspects of abdominal muscle training within the system of global strengthening and flexibility exercises. Figure 5 (Venn diagram) provides a schema of the differences and commonalities of Pilates and core stability training.

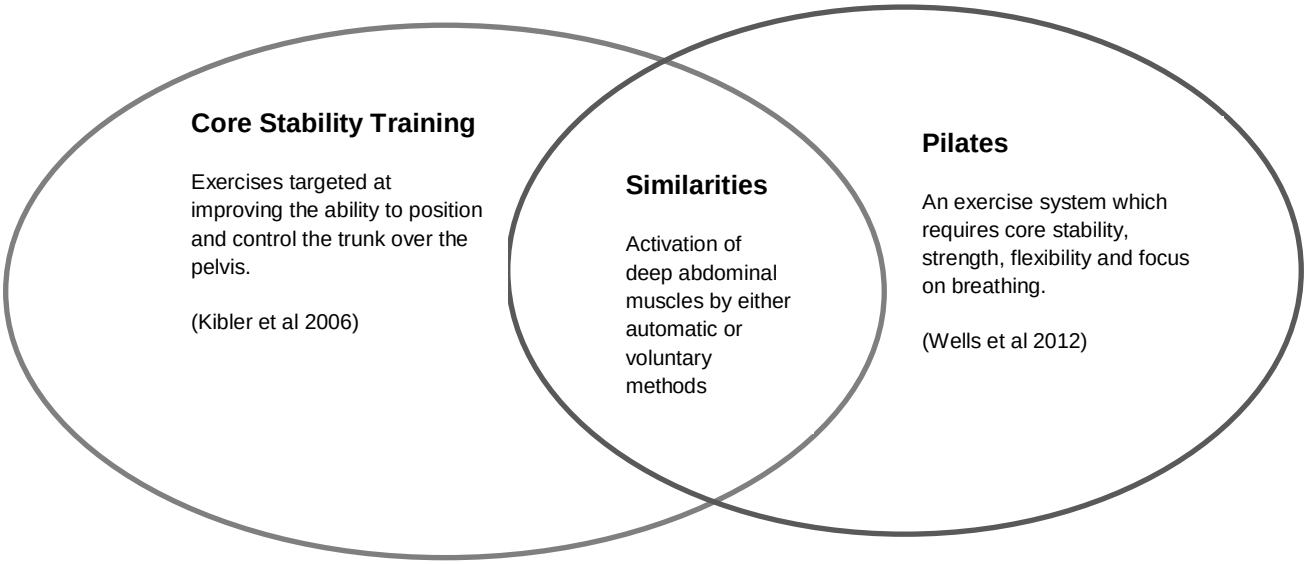


Figure 5: A Venn diagram to highlight overlap between Pilates and core stability training

This literature review chapter aims to critically evaluate the available evidence regarding the effects of Pilates and core stability training upon balance and mobility.

3.2 Literature review

3.2.1 Search strategy

Pilates as an intervention has been widely researched; entering the term 'Pilates' into the search engine 'Embase' generated 329 text results. In contrast the study of Pilates for people with neurological conditions only generated nine text results. The search engines Pubmed', 'Embase' (which includes Ovid Medline and PsycArticles), 'CINAHL' and 'Google Scholar' were searched from 1974- 28th December 2014. To focus the search to meet the specific aims of the thesis, the following search terms were used:-

1) Pilates 'OR' Core Stability 'AND' balance
2) Pilates 'OR' Core Stability 'AND' mobility
3) Pilates 'OR' Core Stability 'AND' walking
4) Pilates 'OR, Core Stability 'AND' Multiple Sclerosis
5) Pilates 'OR' Core Stability 'AND' Stroke
6) Pilates 'OR' Core Stability 'AND' neurological
7) Pilates 'OR' Core Stability 'AND' postural stability .

The results were sorted by relevance and duplications removed. Papers were deemed as relevant if they were published in peer reviewed journals, in English and outcome measures included at least one measure of balance or mobility. In addition a manual search was performed by reading the reference lists of key papers. Papers were included if the study samples comprised healthy people, healthy elderly people or people with neurological conditions. Samples with non-neurological pathologies were excluded (e.g. LBP, HIV, breast cancer and juvenile arthritis); as were peri-natal and sports specific samples. due to the large body of literature which could be deemed not relevant to drawing conclusions regarding the effects of Pilates for people with MS.

3.2.2 Appraisal tools used

The following appraisal tools were used to ensure a systematic and comprehensive critique was undertaken of the papers included in this review:-

- **The PEDro Scale**

The methodological quality of papers was assessed by calculating scores using the Physiotherapy Evidence Database (PEDro) scale, in light of the fact that inadequate quality of clinical trials can distort results (Wood *et al* 2008). Whilst Juni *et al* (2001) suggested that using summary scales for appraising literature may be problematic. They proposed that it is better to evaluate the key methodological components. The use of an appraisal tool, such as the PEDro scale, allows quantification of the quality of research in order to compare methodological quality. This scale has demonstrated to be reliable (Maher *et al*

2003) and valid (Morton 2009) for measuring the quality of research conducted in physical therapy.

- **TIDieR Guidelines**

Appraisal of the reporting of interventions utilised in the studies was performed using the TIDieR Guidelines (Hoffmann *et al* 2014), published to improve the reporting of interventions. The purpose of the TIDier checklist (from which a score can be determined) is to promote detailing interventions in publications such that they could be replicated.

- **AMSTAR**

The methodological quality of systematic reviews was assessed using the AMSTAR measurement tool (Shea *et al* 2007) scores are reported in table 8 . The AMSTAR allows the reader to apply a quantitative approach to the evaluation of the quality of the systematic review. It has demonstrated good reliability, validity and feasibility (Shea *et al* 2009).

- **STROBE**

The STROBE Guideline was used to assistance the critique of the observational studies. A score was not awarded as the intention of the STROBE guidelines was not to be used as an evaluation tool but as guidelines for authors publishing observational studies (von Elm *et al* 2007).

3.2.3 Evaluation of literature

The studies identified in the literature searches are summarised and evaluated in tables 4-7. Table 4 details individual studies evaluating the use of Pilates in healthy people; table 5 details studies which use core stability training as an intervention in healthy people. On reviewing these studies, there was a lack of clarity as to whether the intervention used was Pilates or core stability training. Studies have therefore been categorised, on a pragmatic basis, by the description of the intervention in the title of the article (e.g. 'The effect of Pilates on balance and mobility' or 'The effect of core stability exercise'). Table 6 details studies evaluating the use of Pilates in people with neurological pathology, and table 7 details the one study found which evaluates the use of core stability training in people with neurological pathology.

Author (in order of date)	Design, Intervention and TIDieR score	Sample	Outcome measures	Results	Authors conclusions	PEDro Score	Reason for PEDro Score
Segal et al (2004)	Observational study. Pilates classes 1 x per week for 6 months, taught by 'Stott certified' Pilates instructor. TIDieR score : 10/12	n=47 healthy people, mean age 41, (range 35-48) years. Sample size justified	Finger to floor distance (cm), body composition (lean body mass) and health status (questionnaire) at baseline, 2, 4 and 6 months. Reliability reported.	Significant difference between baseline and follow up at 6 months ($p < 0.01$) for flexibility (median improvement 4.3cm). No significant improvements in body composition or health status.	Participation in Pilates appears to be safe and improves flexibility in healthy subjects.	2/11	Observational study, did not exclude previous or current Pilates involvement, no control group hence no blinding, or randomisation, however thorough reporting of intervention as indicated by TIDieR score.
Johnson et al (2007)	Randomised controlled study. Pilates with certified instructor x 2 per week for 5 weeks. TIDieR score : 6/12	n=40 healthy people (Pilates = 20 control = 20), mean age 27.5 (sd 3.6) years.	Functional reach (FR) test pre and post intervention. Reliability and validity not reported	Significant within group changes for Pilates between pre and post functional reach (FR) ($p=0.01$). Pre FR test =13.61 (sd 2.53) cm, post 14.84 (2.43) cm.	Pilates can improve dynamic balance in healthy people.	5/11*	No blinding of participants or therapists (assessors were blinded), no reporting of between group scores, only within group changes.
Kaesler et al (2007)	Pilot study using Pilates inspired exercise, x 2 per week for 8 weeks TIDieR score : 10/12	n=7 aged 66-71 years. SD not reported.	Postural stability (sway), timed up and go, sit to stand, four scale balance test. Reliability and validity not reported.	Pre –post intervention significant improvements ($p<0.05$) for postural stability and timed up and go.	A short term balance training programme using Pilates inspired exercises may improve postural stability in the elderly.	3/11	Pilot study, no control group hence unable to randomise or blind. Small sample size hence type II error possible.

Table 4: Review of literature of core stability training interventions in healthy people

Reported here: only information relevant to balance and mobility outcome measures, continued over pages 74-78.

Author (in order of date)	Design, Intervention and TIDieR score	Sample	Outcome measures	Results	Authors conclusions	PEDro Score	Reason for PEDro Score
Caldwell et al (2009)	Comparative controlled study. Pilates and 'Taiji quan' exercise classes, control group was an outdoor recreational programme, 15 weeks for 2 x weeks of 50 minutes. TIDieR score : 4/12	n= 127 college students, (Pilates x 51, Taji quan x 35, outdoor recreation x 41), mean age 21.27 (sd 2.24) years. NB: groups differ in size	Strength (dynamometer), balance (single leg stand with eyes closed on a force plate and postural sway assessed) pre and post intervention. Reliability and validity reported.	No increases in strength of lower limbs or balance An effect was found for gender (p=0.001).	Pilates did not affect strength and balance. This sample was already active and fit, which may account for why the effects of exercise classes were negligible.	3/11*	No blinding, no randomisation, no exclusion criteria, similarity of groups not reported at baseline. Difficult to draw conclusions regarding the effects of Pilates since prior prior involvement in Pilates not excluded.
Kloubec et al (2010)	Randomised Controlled Trial 12 weeks of Pilates x 1 hour x 2 per week. Stott Pilates method used. TIDieR score : 7/12	n= 50 healthy people(Pilates x 25 or control x 25), mean 41 (sd 9.12) years). Sample sized justified	Abdominal muscle and upper body endurance, hamstring flexibility posture, balance. Balance assessed using a modified balance board and a counter, recorded each time a participant deviated from mid-point. Reliability and validity not reported.	Within group statistically significant increases in muscle endurance and hamstring flexibility ($p < 0.05$). No significant within group or between group differences for balance and posture	Pilates can improve muscular endurance and flexibility using relatively low intensity Pilates which does not require equipment or a high degree of skill.	6/11	No blinding of assessors or participants. Did not report intention to treat analysis. However did achieve 6/11 which is considered to be the lower limit for rigorous methodology.

Table 4 continued

Author (in order of date)	Design, Intervention and TIDieR score	Sample	Outcome measures	Results	Authors conclusions	PEDro Score	Reason for PEDro Score
Newell et al (2012)	Observational study Pilates classes 1 hour per week for 8 weeks TIDieR score : 8/12	n=9 healthy elderly, mean age 67.8 (sd 5.0).	Inter stride variability and postural sway. Reliability and validity not reported.	Significant within group changes pre and post Pilates for walking speed, step cycle and step length ($p<0.05$).	Pilates may have the potential to improve gait and postural sway in people associated with falls risk.	2/11	Observational study hence no control group. Without control unable to randomize, perform blind assessment or compare groups.
Bird & Fell (2013)	Observational prospective cohort study. This was a follow up to Bird et al 2012 (above) 12 months after Pilates intervention. Once or twice weekly Pilates classes were continued for 12 months. TIDieR score : as for Bird et al 2012 above	n= 30 Older adults (60+),mean age 69 (sd 7) years.Pilates = 15, control = 15. (control were people who declined Pilates)	Medio lateral sway, four square step test, timed up and go (TUG), leg strength Reliability and validity reported.	At 12 months within group changes ($p<0.01$) for medio lateral sway, four square step test and timed up and go and leg strength. Between group significant differences only for leg strength ($p=0.011$).	Pilates may contribute to sustained improvements in falls risk variables. Continued participation for 12 months provided benefits for strength in older adults.	6/11	Lack of blinding of assessors, therapist and participants, similarity of groups at baseline not reported, only 80% of follow up data obtained (as above for Bird 2012).

Table 4 continued

Author (in order of date)	Design, Intervention and TIDieR score	Sample	Outcome measures	Results	Authors conclusions	PEDro Score	Reason for PEDro Score
Mokhtari et al (2013)	Quasi experimental design. Pilates intervention for 12 weeks of Pilates (mat and resistance exercises with bands), 1 x per week. TIDieR score : 3/12	n= 30 (Pilates x 15, control x 15) aged 62-80 years (mean and sd not reported).	Functional reach test, timed up and go. Reliability and validity reported.	Significant differences for Pilates for functional reach (p=0.037). Within group changes: Baseline functional reach 18.19 cm (sd 2.68) week 12, 21.23cm (sd 4.41).	Pilates is efficient at improving balance in the elderly.	3/11	Poor reporting of blinding, numbers and data. Not clear whether changes were within group or between group. Reporting of data was confusing with absence of details which made drawing conclusions difficult.
Pata et al (2013)	Quasi- experimental study. Pilates based exercises intervention, 1 hour x 2 per week TIDieR score : 3/12	n=35 aged 65-87 years , mean 74.4 years (sd not reported) power calculation not reported	Timed up and go, forward functional reach, 180 degree turn Qualitative measures of fear of falling and perception of Pilates Validity and reliability reported for quantitative outcome measures	Significant improvements between baseline and follow up for timed up and go (p<0.001), 180 degree turn (p=0.002), and functional reach (p=0.049).	Pilates may improve balance and mobility and postural instability in older people	4/11	Without control hence unable to randomise, perform blind assessment or compare groups.
Stivala and Hartley (2014)	Single case report. Inpatient rehabilitation with Pilates exercises integrated, 6 days a week for 26 days TIDieR score : 8/12	n=1 (84 year old female with hip fracture and post CVA)	Activities balance confidence scale (ABC), timed up and go, four square step test, forward functional reach, 10 metre timed walk, manual muscle testing of quadriceps and hamstrings. Reliability and validity not reported.	Improvements in gait speed, timed up and go, ABC scale, square step test and muscle strength. Statistical analysis not performed.	The case illustrates the benefits of integrating Pilates into a standard rehabilitation programme and may reduce falls risk	0/11	This was a single case report with no analysis performed. Unable to award any PEDro points.

Table 4 continued

Author (in order of date)	Design, Intervention and TIDieR score	Sample	Outcome measures	Results	Authors conclusions	PEDro Score	Reason for PEDro Score
Hyun et al (2014)	Effect of Pilates vs exercises on an unstable support surface. 40 mins x 3 per week for 12 weeks TIDieR score: 5/12	n= 40 Pilates x 20 vs Unstable support surface exercises x 20). Pilates group aged 70.0 years (sd 2.2) Unstable support surface exercises 69.3years (sd 2.6)	Timed up and go, sway length and speed of centre of foot pressure Reliability and validity not reported	Significant within group (for both groups) decrease in sway length, sway speed and timed up and go ($p<0.05$). For sway speed there was a significant between group difference ($p<0.05$) for unstable support surface	Both Pilates and exercises on an unstable support surface are effective at improving static and dynamic balance, however Pilates may be considered safer.	5/11	No randomisation, blinding of assessors was reported

*In agreement with PEDro score awarded in the systematic review by Cruz Ferreira et al (2011). Other studies in this table were not included in the systematic review by Cruz Ferrira et al

Table 4 continued

Author (in order of date)	Design, Intervention and TIDieR score	Sample	Outcome measures	Results	Authors conclusions	PEDro Score	Reason for PEDro Score
Petrofsky et al (2005)	Observational study. 20 minutes of exercise, 3 days per week for 1 month using a '6 seconds abs machine'. TIDieR score : 5/12	n=13 healthy elderly people mean age 73.1 (sd 7.3) years.	Muscle strength of abdominal and back muscles, balance during functional activities pre and post intervention. Reliability and validity not reported.	Increase in muscle strength in both abdominal flexors and back extensors, increased reach distance and reduced tremor (within group change $p < 0.01$).	Fitness training is beneficial to increasing independence and functional activities of daily living in older individuals.	2/11	Observational study, no control group hence no randomisation or no blinding of assessors or participants, no eligibility criteria stated. Difficult to draw conclusions regarding interventions when eligibility criteria not stated.
Aggarwal et al (2010)	Randomised Controlled Trial Three sessions per week of either core stability or balance training for 40-50 minutes for 6 weeks (core stability exercises detailed in table 7 below, balance exercises in standing to include use of trampoline). TIDieR score : 6/12	n= 30 recreationally active healthy people (core stability x 10, balance training x 10, control x 10) Mean age: core stability group 24.3 (sd 1.6), balance group 25.0 (sd 1.23) and control 24.0 (sd 1.1) years.	Stork balance test and star excursion balance test and single leg hopping stabilisation test pre and post intervention. Reliability and validity not reported.	Within group changes for both exercise groups showed significant improvements for star excursion test and stork tests ($p < 0.05$; none for control group) but not for hopping. Between group changes performed but not reported.	Both core stability training and balance training are effective at improving balance performance	4/11	Blinding of assessors or participants not reported, no between group statistics reported despite being 3 groups, drop outs not reported.

Table 5: Review of literature of core stability training interventions in healthy people

Reported here: only information relevant to balance and mobility outcome measures, continued over pages 79-81.

Author (in order of date)	Design, Intervention and TIDieR score	Sample	Outcome measures	Results	Authors conclusions	PEDro Score	Reason for PEDro Score
Kaji et al (2010)	Observational study. 30 seconds of elbow to toe and hand to heel exercises (plank and reverse plank for 30 seconds). TIDieR score : 5/12	n=17 healthy (young) people.	Centre of pressure during quiet standing with eyes closed before and after 30 seconds of exercises. Reliability and validity reported.	Within group significant decreases in mediolateral sway ($p = 0.0001$), speed of anteroposterior sway ($p = 0.004$), speed of mediolateral sway ($p = 0.004$).	Performing core stability exercises as part of warm-up programs may be useful for temporarily improving postural control during standing in main exercise programmes.	3/11	This was a mechanistic study, taking place on one occasion with no control group hence unable to blind or randomise participants.
Kang et al (2012)	Randomised controlled trial Comparing 30 minutes core stability exercises with control for 8 weeks. TIDieR score : 3/12	n=30 core stability x 15, control x 15), aged 65-80 (mean age and sd not reported).	Berg balance scale, Stability and weight support using force plate analysis. Reliability and validity not reported.	Within group significant changes for Berg balance ($p=0.021$), weight support ($p= 0.014$) and stability ($p=0.003$). Significant between group changes for Berg balance ($p=0.01$), weight support ($p=0.041$) and stability ($p=0.012$)	Core strengthening exercise was effective in improving balance and preventing falls in elderly.	4/11	No blinding of assessors or participants, no mention of randomisation or how participants were allocated to groups or intention to treat analysis.
Hosseini et al (2012)	Three-armed trial. Strength training, core stability training and control, 6 weeks (3 x 1 hour). TIDieR score : 4/12	n= 90 elderly (strength training x 30, core stability x 30 and control x 30), Mean age:strength training 63.3 (sd 4.8), core stability 63.7 (sd 4.2) control 60.7 (sd 5.09).	Y balance test, gait dynamics questionnaire, Strength (bench press and leg press). Reliability and validity not reported.	Within group changes: significant increases in strength of upper ($p=0.003$) and lower limbs ($p=0.004$). Balance improved with both core stability and strength training ($p<0.001$). Core stability training significant differences in gait ($p<0.001$).	'Conducting a period of core stabilisation training improved life independence of geriatric population and will ultimately result in their more contribution to society'	4/11	No blinding of assessor or participants, did not disclose concealment allocation or intention to treat.

Table 5 continued

Author (in order of date)	Design, Intervention and TIDieR score	Sample	Outcome measures	Results	Authors conclusions	PEDro Score	Reason for PEDro Score
Yu & Lee (2012)	Randomised controlled trial 3 x 60-minute Pilates training sessions per week for 8 weeks. Described as core stability training using Pilates. TIDieR score : 7/12	n=40 healthy people (core stability x 20 strength training x 20 control x 20).	Muscle strength as determined by peak torque of knee flexors and extensors. Postural stability measured with Biodex postural stability system. Reliability and validity not reported.	Within group significant increases in core stability group for muscle strength and postural stability ($p<0.05$). Between group significant differences for postural stability ($p<0.05$).	Core stability training using Pilates has a significant effect on lower extremity strength and postural stability in healthy people. Enhanced core stability from Pilates training can prevent musculoskeletal injuries by increasing muscle strength and postural stability thus improving the quality of life.	3/11	Did not report randomisation process, blinding, similarity of groups at baseline, intention to treat analysis, numbers of subjects completing intervention

Table 5 continued

Author (in order of date)	Design, Intervention and TIDieR score	Sample	Outcome measures	Results	Authors conclusions	PEDro Score	Reason for PEDro Score
Freeman et al (2010)	Multicentre series of case studies). 2 x per week for 8 weeks, Individual Pilates based core stability sessions with neuro-therapist. TIDieR score : 10/12	n= 8 people with MS, EDSS 4-6.5, ged 32 -59.	10 metre timed walk, timed up and go, forward and lateral functional reach, MS walking scale -12 , ABC scale, timed single leg stance, visual analogue scale: walking whilst carrying a drink. Reliability and validity reported.	Within group significant difference between pre and post intervention for 10mtw (p=0.019), MSWS-12 (p=0.041, forward and lateral reach (p=0.015 and p=0.012).	The study provides preliminary evidence for the effectiveness of 8 weeks of core stability training for improving balance and mobility in ambulant people with MS.	4/11	Pilot, series of case studies hence no controls or blinding, Unable to compare groups.
Guclu-Gunduz et al (2013)	Randomised Controlled Trial 8 weeks of Pilates vs 8 weeks of abdominal breathing and active extremity exercises (control). Pilates developed by neurotherapist trained in APPI Pilates. TIDieR score : 8/12	n= 26 people with MS (people with MS x 18, control x 8). Age: Pilates 36 (IQR 29-40) control 36(IQR 27.75-45.25) years.	Berg balance scale, timed up and go, muscle strength, ABC scale Reliability and validity not reported.	Within group significant improvement in Pilates for Berg balance scale (p=0.007), timed up and go (p<0.001) and ABC scale (p=0.002). Within group changes for upper extremity and lower extremity strength (p<0.05). No between group comparisons reported.	An 8 week Pilates programme was effective at improving balance, mobility and strength in people with MS.	4/11	No blinding, did not report concealment allocation, intention to treat analysis not reported. Not a PEDro criterion but noteworthy that uneven numbers in intervention and control which questions the rigour of the randomisation and allocation process.

Table 6: Review of literature of Pilates interventions in people with neurological conditions

Continued over pages 82-83

Author (in order of date)	Design, Intervention and TIDieR score	Sample	Outcome measures	Results	Authors conclusions	PEDro Score	Reason for PEDro Score
van der Linden et al (2013)	Feasibility study 12 weeks of Pilates. First 6 weeks 1 hour classes x 2 per week followed by 1 hour per week for 6 weeks supervised group based classes. Mixed methods research, qualitative component. TIDieR score : 10/12	n= 15 people with MS who use a wheel chair (EDSS 7-8) age:51 (sd 8) years. Sample size justified	Sitting stability assessed by sitting functional reach test (centre of pressure), inter scapular distance, visual analogue scale: pain in neck, forced vital capacity, Canadian occupational performance measure, The MS impact scale, fatigue severity scale and qualitative interview. Reliability and validity reported.	Within group significant improvements for centre of pressure (p=0.046), sitting posture (p=0.004), neck pain (p=0.005) and MS impact scale (p=0.006). Enjoyment of class expressed by all.	Pilates appears to be efficacious in improving sitting stability and posture and decreasing pain and is well tolerated.	4/11	Feasibility study hence no control group therefore unable to blind assessors or participants, or compare groups. However high TIDieR score, reporting of validity and reliability of outcome measures and sample size reporting increase the credibility of these findings.
Marandi et al (2013)	Three armed trial 12 weeks of either Pilates or Aquatic (1 hour x 3 per week) vs controls. TIDieR score : 6/12	n=57 females with MS, EDSS less than 4.5, (Pilates x 19 , Aquatic x 19, control x 19), aged 20-40 years .	Six spot step test, timed up and go Reliability and validity not reported.	Significant differences between : Pilates vs control ($p<0.05$), aquatic vs control ($p<0.05$) for adjusted means for both Timed Up and Go and Six Spot Step Test	Both types of exercise had positive effect on dynamic balance compared with control but there were no significant differences between the two types of exercises.	3/11	Did not report concealment allocation, similarities of groups at baseline, blinding of assessors or participants or intention to treat analysis. Sample size calculation was not reported. Larger samples may be required to detect differences between exercise interventions.
Shea & Moriello (2013)	Feasibility study TIDieR score : 9/12	Case report of Pilates for one person with stroke	Lower extremity strength, sit to stand, Berg balance scale, gait speed, stride length, quality of life, thoracic and lumbar curvature. Taken every 3 months for 9 months. Reliability and validity not reported.	Improvements in Berg balance scale, lower extremity strength and quality of life, not posture and gait.	It is feasible to complete a programme of Pilates in conjunction with traditional rehabilitation. It is possible to modify classical Pilates	0/11	Single case study, no analysis performed.

Table 6 continued

Author (in order of date)	Design, Intervention and TIDieR score	Sample	Outcome measure	Results	Authors conclusions	PEDro Score	Reason for PEDro Score
Petrofsky et al (2005)	6 seconds 'abs machine' using resistance bands for exercising abdominal and trunk extensor muscles, 3 days per week for 4 weeks for 20 minutes. TIDieR score : 8/12	n= 14 + 13 control (7 x spinal cord injury, 3 x MS and 4 x stroke).	Computerised posturography during forward reach test and muscle strength and tremor. Reliability and validity not reported.	Within group changes for adominal muscle strength increased by 72% ($p<0.01$), back muscles 62% ($p=0.01$), functional reach ($p<0.01$) forward and lateral. Centre of gravity ($p<0.01$). Between group changes for FR ($p<0.05$).	The 20 minute daily programme only required 20 minutes provided increased function for people with disabilities.	3/11	Did not disclose blinding, randomisation and concealment allocation or intention to treat principles. In addition this study grouped people with different types of neurological pathology i.e. MS and spinal cord injury.

Table 7: Review of literature of core stability interventions in people with neurological conditions

3.3.4 Methodological evaluation of the literature

This section of the literature review will evaluate the methodological rigor and quality of reporting of the articles included in tables 4-7.

Tables 4-7 details the PEDro scale scores for each of the studies, and provides a justification for these scores. In total 20 studies evaluated the use of Pilates or core stability training in healthy people (mean PEDro score 3.3 (range 0-7)). Only five of the studies reached the cut off score of 6/11 (where <6 is an indication of low quality research. Six studies assessed the effects of Pilates and core stability training in people with neurological pathology (mean PEDro score 3 (range 0-4)).

The main areas where these studies lacked methodological rigor was in the lack of blinding of subjects, therapists and assessors. To increase the rigor of these studies, blinding of at least the assessors is essential in order to minimise bias towards the perceived benefits of group allocation, whether intentional or not. Human behaviour is largely affected by belief and hence blinding of assessors is particularly important when outcome measures have a subjective element (Day and Altman 2000). This is supported by the fact that unblinded trials have a tendency towards larger treatment effects than blinded studies (Wood *et al* 2008).

Whilst it is proposed that the blinding of assessors is possible in most circumstances (Wood *et al* 2008), both the blinding of therapists and participants, and the identification of placebo interventions is far more challenging to implement in rehabilitation studies than in drug trials (Day and

Altman 2000). This has the consequence of impeding the credibility of the conclusions drawn in rehabilitation research.

Of the studies included in tables 4-7 none reported the randomisation process with clarity or in detail (with the exception of Bird et al 2012), omitting details such as concealment allocation. –Effective randomisation relies on adequate concealment allocation and it is proposed that concealment allocation is always feasible (Wood *et al* 2008). Mistakes can be made in interpreting data from trials in which randomisation and concealment allocation is not effectively implemented. Studies described as randomised may be assumed to be free of bias; it is possible that this is not always the case. Randomisation which is not computer generated and adequately concealed may be open to deciphering and the effect of the intervention inflated (Schulz and Grimes 2002).

In order to determine the effectiveness of an intervention, such as Pilates training, and whether this has superiority over a control placebo or alternative intervention, such as strength training, it is important to report the treatment effect (the comparison between groups), which should ideally be accompanied by 95% confidence intervals (Bland and Altman 2011). This allows the reader to gain information regarding the estimated effect (Moher *et al* 2010). In this literature review, only nine of the studies reported between group comparisons (Petrofsky *et al* 2005; Johnson *et al* 2007; Caldwell *et al* 2009; Kloubec 2010; Rodrigues *et al* 2010; Irez *et al* 2011; Bird *et al* 2012; Bird & Fell 2013; Marandi *et al* 2013). Two studies included a control group but did not report between group changes (Johnson *et al* 2007; Guclu-Gunduz *et al* 2013). Reporting within-group changes is not sufficient to draw conclusions about treatment, being that the purpose of RCT's is not to determine whether there is an

improvement from baseline but to ascertain the superiority of the intervention above control (Bland and Altman 2011).

The internal validity of studies may be compromised by issues such as attrition bias, which can be reduced by employing intention to treat analysis. Only one study in this review reported intention to treat analysis (Bird *et al* 2012). Studies which exclude participants which do not attend intervention sessions (as in the case of Guclu-Gunduz *et al* (2013)), rather than employ intention to treat analysis, have a tendency to inflate the effect size and overestimate the benefits of treatment (Juni *et al* 2001). In order to reflect a more accurate effect of an intervention such as Pilates, authors could report the numbers of people who did not comply with exercise sessions. This would present a more pragmatic approach to the evaluation of this intervention without biasing towards people who were more motivated to engage in the exercise sessions (Greenhalgh 2008).

Although not specifically a criteria of the PEDro scale, reporting sample size calculations can assist the reader in evaluating the credibility of the study results. In this review only five of the 25 studies, provided sample size calculations (Segal *et al* 2004; Kloubec 2010; Bird *et al* 2012; Bird and Fell 2013 and van der Linden *et al* 2013), therefore it is not possible to know whether studies were adequately powered to draw definitive conclusions. To further add credibility to the results, the magnitude of the clinical effect that the sample size was based on should be included (for example a 20% improvement in walking speed is deemed as clinically significant in people with MS). Studies which are not adequately powered to detect potential between group differences, could

result in a type II error in wrongly accepting the null hypothesis (Whitely and Ball 2002).

When assessing studies which may be underpowered it is worth considering the percentage change which is considered clinically relevant to the target population (Vacha-Haase and Thompson 2004). None of the studies performed in people with neurological conditions in this literature review reported whether changes were clinically significant. In some studies this may be because it has not yet been determined what defines a clinically significant change for the outcome measures (e.g. in studies using the Functional Reach Tests)

In the studies of healthy elderly people Rodrigues *et al* (2010) reported a 0.71 second improvement in 10mtw. Based on their reported data this was equivalent to 9.34% improvement in walking speed. Whilst the effect of Pilates in healthy people cannot be extrapolated to people with MS it is noteworthy that 9.34% improvement is considerably less than the 20% improvement which is considered to be a clinically significant change in walking speed for people with MS (Kragt *et al* 2006).

Finally, drawing conclusions from research depends on the validity and reliability of the outcome measures used to assess the intervention. Fifteen of the journal articles appraised did not report the validity or reliability of the outcome measure used (reported in tables 1-4).

3.4 Systematic reviews

Table 8 details the conclusions drawn from systematic and narrative literature reviews performed assessing the effects of Pilates in healthy people. No

systematic reviews were unearthed relating to Pilates in people with neurological conditions. The 'Assessment of Multiple Systematic Reviews' (AMSTAR) was employed as a tool to critically appraise the methodological quality of the systematic reviews (Shea *et al* 2007). This is an 11 item scale which has good content validity and high reliability for measuring the rigor of systematic reviews in order that the reader can determine the quality of the evidence synthesised (Shea *et al* 2009). Four reviews were included, three of which were described as systematic reviews. Both Cruz-Ferreira *et al* (2011) and Granacher *et al* (2013) scored 10/11 on the AMSTAR rating indicating that the conclusions drawn were based on rigorous methodological protocols, losing one point by not describing publication bias. Cruz Ferreira *et al* concluded that there is strong evidence that Pilates improves dynamic balance in healthy people and Granacher *et al* concluded that Pilates is effective as an adjunct or alternative to balance training in the elderly. Jagannath *et al* (2011) suggested a cut-off score of 4/11 when using AMSTAR; a lower score indicating a poor quality systematic review. Wells *et al* (2012) was awarded 3/11 however it is noted that the purpose of the review by Wells was to define Pilates rather than to evaluate this as an intervention.

Author	Type of review and population	Authors Conclusions	AMSTAR score
Bernardo (2007)	Narrative review of healthy adults. Used Pilates as a search word, 277 articles generated; 39 were in peer reviewed journals (others were published in magazines and newspapers). All were observational studies or uncontrolled experimental studies. Only 3 were performed in healthy adults.	Cautious support for effectiveness of Pilates in healthy adults for improving flexibility, TrAb activation and lumbar pelvic stability. Caution due to small sample sizes and poor experimental design. Well-designed experimental studies that randomize subjects, utilise a control group, clearly define Pilates method (including skill of execution of exercises), calculate statistical power and use valid and reliable methods to measure outcomes would contribute to a body of scientific evidence for Pilates efficacy.	6/11
Cruz-Ferreira et al (2011)	Systematic review of healthy adults. Pilates used as search word, 16 studies met criteria, research assessed using PEDro scale.	PEDro scores ranged from 3-7 (mean 4.1) indicating low scientific rigor. Conclusions suggest that there is evidence that Pilates increases flexibility, dynamic balance and stabilisation of core posture. No evidence for postural alignment, strength and static balance.	10/11 (1 point lost for not reporting publication bias)
Wells et al (2012)	A systematic review to define Pilates, using the search term 'Pilates'. 2182 papers generated of which 119 fulfilled criteria	Based on this systematic review the definition of Pilates is 'a mind body exercise approach requiring core stability, strength, flexibility and attention to muscle control and breath. Exercises may be floor based and include specialised equipment'. None of the papers reviewed made reference to the traditional principles of Pilates.	3/11 (This review was intended to define Pilates as opposed to evaluate interventions)
Granacher et al (2013)	Systematic review assessing the effects of Pilates and core stability training upon the trunk muscle strength, balance and falls of seniors (>65 years). Nine studies met the inclusion criteria of using core stability training or Pilates upon trunk muscle strength, functional performance and falls of older people.	Pilates and/ or core stability training can be used as an adjunct or even an alternative to traditional balance and/ or resistance programmes for older adults. Pilates exercises are easy to administer and require little space and equipment.	10/11 (1 point lost for not reporting publication bias)

Table 8: Systematic and literature reviews of Pilates in healthy people

3.5 Critique of individual studies using Pilates or core stability training for improving balance and mobility in people with neurological conditions

The aim of this chapter was to determine from the existing literature the effects of Pilates upon the balance and mobility of people with MS. This section will focus specifically on the research performed using Pilates or core stability training in people with neurological conditions.

This section expands on the details and further critiques the methodology and findings of the studies reported in tables 6 and 7. The details of each study include intervention, sample (and sample size calculation), outcome measures (including their reliability and validity), results (including statistical significance) and PEDro scores are included in tables 6 and 7.

Freeman *et al* (2010) performed a replicated series of single case studies to undertake a preliminary exploration of the effect of Pilates; this consequently served as a basis for designing our clinical trial. The aims of the pilot study were to explore the effectiveness of a programme of core stability training in the target population, and to determine which outcome measures were the most responsive in capturing any changes that occurred. Due to the nature of the study (single case studies), it is only possible to conclude that an 8 weeks course of Pilates based core stability training may result in improvements in walking (10mtw and MSWS-12) and balance (forwards and lateral functional reach) in ambulant people with MS.

The study performed by Guclu-Gunduz *et al* (2013) compared Pilates with a control, which consisted of abdominal breathing exercises and active extremity exercises thus omitting the targeted voluntary activation of the deep abdominal

muscles performed during Pilates. Within group improvements were reported for the Berg Balance Scale, Timed Up and Go test, ABC scale and upper and lower extremity strength. It is noteworthy that the groups were unevenly matched (intervention n=18, control n=8). The internal validity of studies may be compromised by issues such as selection bias (bias in the allocation of groups) resulting in uneven group sizes. This could have been a result of attrition bias (people dropping out of the control group), however this was not reported by the authors. As previously discussed (page 87), employing intention to treat principles to the analysis can help accurate reporting and enhance meaningful interpretation of the effects of an intervention (Juni *et al* 2001), but this was not undertaken.

In the study by Guclu-Gunduz *et al* (2013) the Pilates intervention group exercised for one hour twice a week, whereas the duration of exercise for the control group was not stipulated. The Pilates intervention group performed exercises in supine, quadruped, sitting on the gym ball and standing, hence there may have been a task specific component, to the Pilates intervention, in that the outcomes measures were the BBS, TUG and ABC scale. It was not detailed which specific exercises the control group performed. Hence, it is not possible to determine whether the improvements in balance and mobility seen in the Pilates group could be attributed to the voluntary activation of the deep abdominal muscles, the intensity and duration of exercises performed or the task specific nature of the standing balance exercises.

Marandi *et al* (2013) undertook a three arm study comparing Pilates, aquatic exercise, and a control (no details of control reported). Pilates and aquatic exercise both resulted in significant improvements in balance in comparison to

control as measured by the Six Spot Step Test. On comparing the Pilates and aquatic therapy the differences were not significant (p=0.95) for balance. This may be because Pilates is not superior to aquatic exercise in improving balance, or due to the sample size, which may not have been adequately powered to detect between group differences in exercise interventions. This is currently unclear as sample size/ power calculations were not reported. As the between group mean differences were small it is probable that Pilates is not superior to aquatic exercises.

The feasibility study performed by van der Linden *et al* (2013) used mixed methods research to investigate the effects of Pilates for people with MS who were wheel chair dependent (EDSS 7-8). A strength of this study is that the Pilates intervention was designed by experienced MS specialist therapists and delivered by Pilates instructors trained in working with people with neurological conditions. Within group improvements on objective measures were reported for sitting balance, posture and pain. The qualitative data revealed that participants all enjoyed the classes, reporting their experience as 'overwhelmingly positive'; this was reflected by the high adherence rate of 81% over six weeks. Longer term adherence was not assessed. Long term participation in exercise has been considered a necessity for maintaining performance of activities of daily living and quality of life for people with MS (Rietberg *et al* 2004); enjoyment of exercise serves to improve adherence (Hale *et al* 2012). To date longer term adherence to exercise regimes has not been established in MS.

Petrofsky *et al* (2005) described the intervention used in their study as core stability training using a six seconds abs machine with resistance bands for trunk muscle strengthening. This intervention did not employ voluntary

recruitment of the deep abdominal muscles, however it could be assumed (although it was not measured) that there was automatic recruitment due to destabilisation of the spine (Hu *et al* 2012). The phenomenon of automatic recruitment of the deep abdominal muscles is discussed in full on page 250. The study sample was heterogeneous, comprising people with a variety of neurological conditions (MS, stroke and spinal cord injury); some of whom were paraplegic while others were ambulant. The control group were healthy people. A more homogenous sample, with controls from the same population is necessary establish effectiveness of this intervention. Hence the conclusions drawn by the authors that “functional reach significantly improved in people with neurological conditions with trunk muscle strengthening” should be viewed with considerable caution.

3.6 Diversity in Pilates and core stability training as an intervention

In comparing studies investigating Pilates it is noteworthy that they differ greatly in the type, intensity and frequency of the exercises which constitute ‘Pilates’ making comparison of outcomes difficult. Table 9 describes examples of the differences in Pilates interventions and table 10 highlights the exercises used in core stability training. The lack of consistency in the definitions and delivery of Pilates interventions has been reported by Bernardo (2007). To explore this issue in more depth, the TIDieR guideline for appraising the reporting of interventions, were utilised in this literature review. The TIDier guideline does not employ a numerical cut off point to categorise the research according to quality, however it does highlight publications which do not detail interventions in sufficient detail for replication. Studies included in this literature review which

scored less than 4/12 (Kang *et al* 2012; Mokhtari *et al* 2013; Hosseini *et al* 2012) failed to include details such as where and who provided the intervention, whether any modifications had been performed and the adherence to intervention. Higher scoring studies, such as Freeman *et al* (2010) and van der Linden *et al* (2013) reported details such as qualifications of persons providing the interventions, any assistance from support staff, location of the delivery of interventions and methods used to records adherence (e.g. a tick box diary).

Pilates exercises have diversified as they have been used to accommodate the different needs of client/ patient populations, and have evolved in line with current evidence (Wells *et al* 2012). In addition, the removal of trademark restrictions over the term 'Pilates' has resulted in dilution of the original techniques and widespread alteration (Brown, 2002). The traditional Pilates principles of concentration, centering, control and flow (described in table 3 introduction section) were not reported in 92% of studies included in the systematic review by Wells *et al* (2012). It is possible that this indicates that a less traditional approach to Pilates is being used in clinical populations.

Whilst it would seem appropriate for some of the exercises from the original Pilates repertoire to be viewed as unsuitable (and even aggravating) for people with conditions such as back pain. Furthermore it may not be possible for people with MS to perform Pilates at the intensity detailed in studies with younger healthy populations due to the nature of the disease giving rise to fatigue.

Author	Format of Pilates	Instructor training	Details of Pilates intervention
Johnson et al (2007)	Group sessions (5 per group)	Certified Instructor	All exercises performed on the reformer using springs and bands for resistance with arms and legs; limited details of exercise provided. No details of duration of each session. No mention of voluntary recruitment of deep abdominals or stretching.
Caldwell et al (2009)	Group sessions (participants per group not reported)	Trained Instructor	Two x 50 minute sessions per week. Authors did not disclose any information regarding the type of exercises performed, such as position of exercise (standing or on the mat) , or the use of resistance bands or reformer. Stretching, breathing exercises or relaxation not disclosed.
Rodrigues et al (2010)	Small group sessions (participants per group not reported).	Physical therapist certified in Pilates method	Detailed reporting of Pilates exercise intervention which included 10 minutes of global stretching, 40 minutes of conditioning exercises and 10 minutes of relaxation. Exercise performed with reformer and resistance bands, and in standing, supine and prone.
Bird et al (2012)	Group session with 6 people per session.	'Pilates alliance' trained instructor	60 minutes twice per week using reformer and mat work to include standing balance exercises. Home exercises given. Not disclosed whether resistance bands used.
Newell et al (2012)	Small group session (participants per group not reported).	Qualified instructor	8 weeks of core stability addressed by 'abdominal bracing and pelvic tilts'. Theraband for resistance, Swiss ball, weights and wobble board used. Exercises undertaken in supine. Noteworthy that Pilates training with APPI and DMA clinical Pilates opposes abdominal bracing as they propose it is counterproductive to increase spinal stiff to the degree achieved with abdominal bracing (Withers, www.ausphysio.com/Files/files-filename-24.pdf , personal communication with DMA clinical Pilates, 2012)

Table 9: Examples of Pilates interventions used in research

Continued overleaf

Author	Format of Pilates	Instructor training	Details of Pilates intervention
Cruz-Ferreira et al (2013)	Group sessions (participants per group not reported).	Pilates certified instructor following the 'Body Control' method.	60 minutes x 2 sessions per week for 6 months. Participants learned 34 exercises in standing, supine, prone and 4 point kneeling using resistance bands and weights. Focus on alignment, breathing, stretching and both lumbo-pelvic and scapulo-thoracic stability.
Mokhtari et al (2013)	Group sessions (participants per group not reported).	Therapist or instructor training in Pilates not disclosed.	No details of exercises, no discussion about exercises used. No mention of voluntary recruitment of deep abdominal muscles. Bands used for resistance, exercises performed on mat. No disclosure as to whether exercises were standing or supine/ prone/ 4 point kneel.
van der Linden et al (2013)	Group sessions (participants per group not reported).	Qualified Pilates instructor, exercises selected by MS specialist physiotherapist and Pilates instructor.	Focus on engaging core muscles, to include reaching and passing a ball and weights and theraband for resistance. Exercises performed in seated as wheelchair dependent population.
Note: none of these studies performed one to one sessions			

Table 9 continued

Author	Description of core stability intervention
Petrofsky et al (2005)	The 6 second 'abs machine' was used. Subjects leant against rubber resistance bands which aimed to strengthen rectus abdominis, external oblique, spinal extensors and deep abdominal muscles; used in sitting. 10 minutes of passive stretching prior to use, 20 minute exercise sessions, 3 days per week for 1 month.
Aggarwal et al (2010)	Focus on voluntary activation of TrAb and lumbar multifidus, 4 point kneel with arm extension, seated medicine ball rotation, seated and squats using Swiss ball, lunges, oblique pulleys, planks using Swiss ball, bridges using Swiss ball

Table 10: Details of core stability training interventions used in research

3.6 Outcome measures used in Pilates and core stability studies

Twenty one of the studies included in this chapter assessed balance as an outcome measure, and 11 assessed mobility (for a detailed description of ICF definitions of balance, mobility and walking refer to page 28-30). A possible reason for this is that researchers did not anticipate that Pilates or core stability training would impact on mobility as much as balance. This is perhaps unsurprising given the focus that Pilates places upon training the trunk muscles, and the knowledge the trunk is associated with balance performance both in healthy people (Suri *et al* 2009) and those with MS (Lanzetta *et al* 2004).

Various methods of measuring balance have been used in the studies ranging from laboratory based measures of postural sway (Kaji *et al* 2010) to clinician rated measures of function (such as the Functional Reach Test, Johnson *et al* 2007). Improvements in these measures have been noted in studies evaluating Pilates interventions (Tables 4-7).

The effects of Pilates and/or core stability training upon mobility are less well documented. Nine studies used a timed up and go (TUG) test as a measure of mobility (not walking), four of these were in neurological populations, two in elderly populations (Freeman *et al* 2010; Bird *et al* 2012; Bird & Fell 2013; Guclu-Gunduz *et al* 2013; Mokhtari *et al* 2013; Stivala & Hartley 2014). Four studies used the 10mtw to measure walking (Freeman *et al* 2010; Rodrigues *et al* 2010; Shea & Moriello 2013; Stivala and Hartley 2014) (three were in neurological samples). Only the studies comprising aging or clinical populations (people with MS and stroke) included measures of mobility, presumably because healthy people are unlikely to experience mobility impairments.

In summary, to date balance measures have been frequently used to assess the impact of Pilates and core stability interventions whereas mobility measures have not.

3.4 Summary of literature review evaluating Pilates and/or core stability interventions

Based on current evidence both Pilates and core stability training appear to positively influence balance in both healthy people and those with neurological pathology. However these conclusions are drawn tentatively due to the poor rigor implemented both in the methodological design of studies and the reporting of research. Pilates and core stability training do not appear to have superiority over other forms of exercise (such as strength training and aquatic exercise) in improving balance, although studies to date have not been powered to determine comparative effectiveness. The impact on balance has been investigated more frequently than mobility in the studies evaluated. There is some preliminary evidence to suggest that Pilates may improve mobility in elderly people and people with MS.

There were no reported ill effects or harms in any of the studies reviewed. It is reasonable to suggest therefore that Pilates is likely to be a safe form of exercise for healthy people, the elderly and more tentatively for people with neurological conditions, including MS.

The paucity of high quality research in this field highlights the requirement for well-designed adequately powered RCT's to enable evidenced based conclusions to be drawn as to the effects of Pilates on balance and mobility on people with MS, and to confirm the safety of this exercise.

Section one, Chapter Four Literature review of outcome measures

4.1 Introduction to outcome measures

To draw meaningful conclusions regarding the effectiveness of interventions, the clinical appropriateness and scientific rigor of the outcome measures used must be considered (Cohen *et al* 2012). No single outcome measure is able to capture a reflection of all changes in all populations (Cohen *et al* 2012), with a general consensus that there is no single ideal outcome measure (Amato and Portaccio 2007). Some have proposed an urgent requirement for a core set of outcome measures for evaluating the effects of exercise in MS (Rietberg *et al* 2004), and work to establish these has recently been undertaken (Paul *et al* 2014). This chapter focuses upon the rationale for choosing the measures used in this clinical trial, and provides a critical appraisal of these measures. It is noteworthy that since the initial design of the RCT there has been a proliferation of literature regarding the use of many of these outcome measures, especially those used to monitor walking in MS.

The aim of this clinical trial was to evaluate the effectiveness of Pilates exercises for improving balance and mobility in moderately disabled people with MS. The outcome measures were chosen on the basis of pilot research, which comprised a series of multicentre single case studies (Freeman *et al* 2010), described in detail on page 82.

4.2 Psychometric properties: Definitions

Three scientific properties are important in determining the usefulness of an outcome measure: reliability, validity and responsiveness (Hobart *et al* 1996). This section will define and discuss each of these.

4.2.1 Reliability

Reliability: the ability to produce results that are accurate, consistent and reproducible (Field 2009). Internal consistency, test-re-test reliability and rater reliability fall under the umbrella term of reliability (Finch *et al* 2002).

Internal consistency: determines whether several items that propose to measure the same general construct produce similar scores. This is usually measured with Cronbach's alpha coefficient, which is calculated from the pairwise correlations between items. It is widely accepted that Cronbach's alpha should exceed 0.70 (Lohr 2002).

Test-retest reliability: measures the stability of an instrument over time. It is assessed by undertaking the 'test' on the same group of subjects on different occasions, and determining the correlation between scores (Hobart *et al* 1996).

Rater-reliability: measures the agreement between assessors (inter-rater), by correlating scores of two (or more) assessors on one occasion, or within assessor ratings over two (or more) occasions (intra-rater). It is most commonly defined by an intra-class correlation coefficient (ICC), which is ideally reported with 95% confidence intervals to reflect where the true correlation lies within the population sampled (Hobart *et al* 1996).

4.2.2 Validity

Validity: determines whether the instrument measures the concept that it is \ intended to measure and can be categorised by content, criterion related validity and construct validity (Field 2009).

Content validity: is concerned with whether a measure appears to be measuring what it intends to measure. Furthermore it indicates that a measure is composed of a comprehensive sample that completely assesses the construct of interest (Finch *et al* 2002).

Criterion related validity: Criterion validity measures the test against other validated tests (referred to as the "gold standard") of the same construct. For example a laboratory based measure of postural sway could be considered the gold standard measure for balance against which the Functional Reach Test would be validated. This is described as concurrent if the measures are taken at the same time (on the same occasion) (Greenhalgh *et al* 1998).

Specificity and sensitivity are defined as a special form of validity relating to binary measures which produce information about the diagnostic accuracy of a test. They may be considered a subset of criterion validity. Sensitivity refers to a test's ability to identify the presence of a condition correctly. Specificity refers to the test's ability to exclude the presence of a condition correctly (Greenhalgh 2008).

Construct validity: indicates that a test measures the concept it is theoretically predicted to measure. In the absence of a gold standard construct validity can be applied. This involves forming theories about the attribute and then

assessing whether the results are consistent with the theories (Finch *et al* 2002).

Convergent and discriminant validity: are the two subtypes of validity that make up construct validity. Convergent validity refers to the degree to which two measures of constructs that theoretically should be related, are in fact related. In contrast discriminant validity tests whether concepts or measurements that are supposed to be unrelated are, in fact, unrelated (Finch *et al* 2002).

4.2.3 Responsiveness

Responsiveness defines an instrument's ability to measure change over time (Guyatt et al 1987; Baert et al 2014). Currently there is no universally accepted consensus as to the best method to determine responsiveness (Kieseier and Pozzilli 2012), and differing methods are available which include distribution based methods such as effect sizes and standard error of measurement (SEM) (Tyson and Connell 2009), and anchor based methods; each with advantages and disadvantages (Man-Son-Hing *et al* 2002).

When evaluating the measurement properties of the outcome measures throughout this chapter guidance was sought from Hobart et al (1996), Greenhalgh et al (1998), Greenhalgh (2008), Field (2009) and (Finch *et al* 2002).

4.3 Measuring walking

Walking impairment is a major concern for people with MS (Heesen *et al* 2008), with significant social and economic implications (Pike *et al* 2012). It is a key determinant of quality of life (Yildiz 2012). This clinical trial measured the effects

of Pilates upon ambulant people with MS, on mobility (specifically walking) and balance. In order to provide a rounded reflection of this, assessment was undertaken both from objective clinician rated assessments, and from the perspective of the person with MS. This literature review will focus more heavily on the 10mtw as this was the primary outcome measure for which the statistical power of the study was based.

Measuring walking time and calculated speed (velocity) is one method for quantifying walking impairment. Its importance is underlined by research demonstrating the strong relationship it has with important activities such as community ambulation (Kempen *et al* 2011; areas under the ROC curves 0.74 - 0.86, with small 95% confidence intervals). Further, a study by Yildz *et al* (2012) highlighted that up to 53% of their sample of 605 people with MS across four countries reported avoiding ADL due to concerns about impaired walking speed (no correlation reported). Using walking tests has been recommended by a number of authors as an effective means of evaluating walking speed within clinical settings (Gijbels *et al* 2012; Kieseier and Pozzilli 2012b; Yildiz 2012).

4.3.1 The primary outcome measure: 10 metre timed walk (10mtw) and calculated walking velocity

There is extensive published data regarding the psychometric properties of the 10mtw, much of which has been published since the conception of this clinical trial.

Description: The 10mtw is a timed walk over a set distance (10 metres), either at a self-selected speed or fastest speed, using a person's usual walking aid

and/ or orthotics (Kieseier and Pozzilli 2012). This can be measured in time (seconds) or speed/ velocity (distance divided by time). In the literature speed and velocity are used interchangeably with the same meaning. Walking velocity is calculated by dividing the distance (10 metres) by the number of seconds taken to walk this distance. Transforming time taken to walk 10 metres in seconds to velocity (metres per second) creates more normally distributed data which is better suited to the assumptions of testing parametric data as results based on velocity rather than time are less likely to be influenced by skewed distributions (Hobart *et al* 2013). A critical review of the timed 25 foot walk test (T25fwt = 7.62metres) will be incorporated within this section as this is commonly used internationally as a 10mtw equivalent.

Purpose: to assess walking speed in metres per second as a measure of walking impairment over a short distance (http://www.rehabmeasures.org/Lists/RehabMeasures/DispForm.aspx?ID=901, access 19th October 2014; 14:26h).

ICF domain: activity

4.3.2 Psychometric properties

A systematic review by Tyson and Connell (2009a) recommends the use of the 10mtw in people with neurological conditions, providing robust evidence from an array of studies that have consistently demonstrated it to be psychometrically sound across a range of conditions and abilities.

Feasibility

Short timed walking tests (10mtw, T25fwt) are widely considered as being of high practical value in the clinical setting, requiring little time, space or equipment (Kieseier and Pozzilli 2012) .

Reliability

High test-retest, inter-rater and intra-rater reliability of the short timed walk tests has been demonstrated, both when measured multiple times on the same day and at one week intervals over 3-4 weeks (see table 11) . This is the case both for mildly (EDSS < 4) and moderately disabled individuals (EDSS > 4) (Kieseier and Pozzilli 2012). Due to the nature of neurological fatigue it is plausible that the time of day could affect the reliability of short walking tests (Burschka *et al* 2012), however, a multi-centre trial demonstrated that the 10mtw was unaffected by time of day despite changes in subjective fatigue (Feys *et al* 2012). These results were mirrored by Morris (2002). Consequently the time of day when performing the 10mtw does not appear to need to be consistent in order to be reliable.

Author **(date order)**	**Test**	**Sample**	**ICC**	**95% CI**	**SEM**	**MDC**
Paltamaa *et al* (2005)	10 metre timed walk (normal speed)	n= 19 ambulant people with MS	0.93	0.72-0.98	0.10m/s	Not reported
Kieseier & Pozzilli (2012)	10 metre timed walk (normal speed)	Literature review of outcome measures in MS	Test re-test at 1 week: 0.91	0.81-0.96	0.09 m/s	Not reported
Learmonth *et al* (2013)	Timed 25ft walk (not reported whether normal or fastest speed)	n=82 EDSS 3.5	0.99	0.98-.099	1 second	2.7 seconds (equivalent to 36%)
Learmonth *et al* (2013)	Velocity calculated from Timed 25ft walk (feet/second)	n=82 EDSS 3.5	0.99	0.98-.099	0.1 feet/ seconds	0.1 feet/ second (equivalent to 36%)
ICC= intra-class correlation co-efficient, CI= confidence intervals, SEM=standard error of measurement, MDC=minimal detectable change						

Table 11: Published reliability statistics for short timed walking tests

Validity

The validity of the 10mtw is strongly supported for use in people with MS, as highlighted below:

Moderate to strong correlations have been reported between the 10mtw and an array of measures known to measure similar constructs. For example a strong relationship was reported between the 10mtw and the Modified Functional Walking Categories (n=156, r =0.74-0.86) (Kempen *et al* 2011). Dalgas et al (2012) also reported strong correlations between the 10mtw and 6- minute walk test (r=0.95) walking at fastest speed (mean EDSS 3.8). Dalgas *et al* surmised that this may be because walking capacity is determined by neural impairments regardless of the walking test distance.

Furthermore walking speed over short distances such as 10m and 25ft have been found to correlate moderately to mean daily stride count (r=0.58) (Gijbels et al 2010). The relationship between walking tests and disability level reported by Kieseier and Pozzilli (2012) are detailed in table 12. The use of the ICC to determine validity, as has been undertaken by Kieseier and Pozzilli (2012) has been criticised by authors such as Zaki *et al* (2012), who highlight that the ICC is affected by the data range; if variance between scores is high, the ICC will also be high.

The validity of short walking tests has been further evaluated (and supported) against an established patient reported measure: the MSWS-12. A number of studies, including Hobart *et al* (2003) and McGuigan & Hutchinson (2004), have reported moderate correlations between timed walking tests and the MSWS-12 (this is expanded upon in section 4.3.2 below).

Slower walking times on the 10tw are an important predictor of perceived difficulties in self-care 1(n=120 people with MS) (Paltamaa *et al* 2007). Furthermore a one metre per second change in the 10mtw had good sensitivity and specificity for predicting limitations in ADL's (Kierkegaard *et al* 2012). The 10mtw has been found to discriminate between pwMS and healthy controls and also to differentiate between mild and moderate levels of disability (p=0.01) (Kieseier and Pozzilli 2012).

Criterion validity **reported in Kieseier and Pozzilli (2012)**	**10 metre timed walk** **(usual speed)**
30 metre walk test	ICC 0.85 (95% CI 0.74-0.92)
Six minute timed walk	$r^2 = 0.80$ ($p < 0.01$)
EDSS	r = 0.69
EDSS ≤4	ICC 0.70 (95% CI 0.42-0.86)
EDSS ≥4	ICC 0.85 (95% CI 0.66-0.94)
ICC= intra-class correlation co-efficient, CI= confidence intervals, EDSS= Extended Disability Status Scale	

Table 12: Criterion validity of the 10 metre timed walk

Self-selected or fastest walking speed

The 10mtw can be performed at either a self-selected speed or at fastest speed (Dalgas *et al* 2012) and is also affected by other factors which include both dynamic or static start (Gijbels *et al* 2012), hence these require standardisation and accurate documentation. Self-selected speed has been frequently used both in MS studies (Morris 2002; Nilsagard *et al* 2007;Barrett *et al* 2009); and in routine clinical practice.

In recent years however, subsequent to the implementation of this clinical trial, research suggests that fastest speed may be the method of choice for reflecting walking capacity (Gijbels *et al* 2012).

Responsiveness

Walking velocity has been reported to be a highly responsive measure of walking impairment, for example as demonstrated in a sample of 120 people with MS (EDSS 0-6.5). This was calculated using both anchor and distribution methods, at a time period of one year apart (Paltamaa *et al* 2008).

Few MS studies have explored the responsiveness of walking tests. Those performed to determine the responsiveness of walking tests have been in samples in which walking deteriorated, rather than in those where improvement could be expected (for example after physical rehabilitation) (Freeman *et al* 2013). Consequently Baert et al (2014) performed a study to establish their responsiveness in this context in a multi-centre sample of 290 people with MS. They concluded that longer walking tests, such as the two and six minute walk tests, and self-report measures (MSWS-12) are more responsive in detecting change after physical rehabilitation than the 10mtw.

Outcome measures are subject to both floor and ceiling effects which can reduce responsiveness. Bethoux and Bennett (2011) reported a floor effect on the T25fwt, potentially making it less responsive in people with mild disability (< EDSS 3.5). Bearing in mind the target population for this clinical trial was individuals who were moderately disabled, the effect may be less prominent.

Clinically significant change

It is considered that a 20% change in walking speed is clinically significant in ambulant people with MS (Kragt *et al* 2006; Hobart *et al* 2013; Learmonth *et al* 2013). Initially the clinical impact of a 20% worsening in T25ftwt was documented by Kragt et al (2006). However this was based on comparing scores of people who experienced at least a 20% change in the T25fwt, with the nine hole peg test and the Guys Neurological Disability Scale (GNDS). Weak but significant correlations between T25ftwt and GNDS change scores over this time period were reported ($r=0.23$, $p<0.05$). The relevance that the correlation to mobility is questionable and may not be the most appropriate measure to define clinically significant change in walking.

Clinically significant changes in the 10mtw have been reported as an increase in walking time (seconds) of 23% and a deterioration of 30% (Nilsagard *et al* 2007). Other studies have reported a change in velocity of either an increase of 0.17m/s or decrease of 0.12m/s as being clinically significant (Morris 2002). On the basis of these studies it can be concluded that the clinically significant change of the 10mtw lies between 20- 30% change for moderately disabled people with MS.

Minimally Clinical Important Differences (MCID) of short walking tests

Kieseier and Pozzilli (2012) reported that there is little consensus amongst MS researchers as to the MICD for the 10mtw, although attempts to define this for walking speed have been reported. Percentage improvements have been derived by observing the intra-patient variability in clinically stable patients over a set distance. This information however does not consider the direction of

change in walking speed, therefore represents a value from which patients are considered not to vary over brief periods, not the extent to which improvement can be compared with deterioration (Coleman *et al* 2012).

On the basis of T25ftwt data from Fampridine trials, Coleman et al (2012) suggested that the MICD for walking speed is a 17.2% relative improvement from baseline scores. Further analysis suggested that the MCID was smaller in patients with faster baseline walking speeds and larger in those with slower baseline speeds. It is not unreasonable to suggest that these values may also be similar for the 10mtw.

4.3.3 Discussion points

Variability in clinically significant changes

When using a short walking test to evaluate changes in walking, speed variability should be considered. A 20% change is considered to be the threshold that indicates a reliable change in the T25fwt (Schwid *et al* 2002). However within day variability of > 20% has been reported when baseline walking velocity was <1.2m/s (Feys *et al* 2014). In contrast, using pooled data from Fampridine drug trials (n = 533 ambulant people with MS), Hobart et al (2013) determined that the variability of speed in the T25fwt was small. Within and between visit averages ranged from 7.2% to 16.3%. In summary, whilst there may be variability between days, over a longer time frame variability appears to be lower than 20% in ambulant people with MS.

Community ambulation

Community walking is important as it encompasses the participation domain of the ICF. Although gait speed and ability to walk in a community setting are related, they are however not the same things and require different outcome measures (Kempen *et al* 2011). It has been demonstrated that people with unlimited community walking ability have a minimal gait speed of 1.63 metres/ second compared to 0.48 metres/second in those limited to walking inside the house (Kempen *et al* 2011). One method of measuring community ambulation is accelerometry; which has demonstrated to be reliable and valid in MS (Learmonth *et al* 2013). Accelerometery could have been a viable outcome measure for this clinical trial. The trial, however, was intended to be pragmatic and hence, given that the use of accelerometry was not routinely used within physiotherapy practice, accelerometry was not chosen as a measure. Potential outcome measures not used in this clinical trial are fully discussed in the discussion of the clinical trial page 178.

Measuring biomechanical changes in gait characteristics

Measuring walking speed alone does not capture the biomechanical changes in gait which may occur as a result of impairments caused by MS. Morris (2002) reported that people with MS, in comparison to matched controls, not only walked more slowly but also had reduced stride length and twice as much variability in gait performance. These findings were supported by Martin et al (2006) who found that even people mildly affected (EDSS <2.5) walked with reduced speed and stride length; and also with altered ankle muscle recruitment. Socie et al (2013) further confirmed these findings, again demonstrating greater variability in step length and step time in people with MS

than matched controls. Biomechanical assessment of gait is beyond the scope of this chapter however it is worth noting when designing clinical trials that walking speed alone does not capture all walking impairments. In this trial we could have chosen a valid biomechanical assessment of walking, however as this multi-centre trial was designed to be pragmatic and reflective of clinical practice, sophisticated and expensive measures of biomechanical assessment were not chosen.

4.3.4 Summary of 10mtw

The 10mtw is a clinician rated measure of activity which has demonstrated good reliability, validity and responsiveness in a range of different MS samples. It is a cheap and easy measure to implement and is routinely used in clinical practice to objectively measure walking speed. In more recent years, since the conception of this clinical trial, some authors have suggested using longer tests, such as the two or six minute walking distance tests since there is evidence to suggest they may better reflect walking ability and be more responsive to changes occurring with rehabilitation interventions (such as exercise), in a moderately disabled population. As the trial was designed to be pragmatic, based on both the information gained from the single case study pilot research and the available literature at that time, the 10mtw at a self-selected speed was chosen as the primary outcome. It was considered to be a safe, valid, reliable and responsive measure which would be feasible to administer in a multi-centre trial implemented within an NHS setting (Freeman *et al* 2010).

4.4 Secondary outcome measures

4.4.1 12 item-Multiple Sclerosis Walking Scale (MSWS-12)

Description: This 12-item self-report questionnaire was formulated to evaluate the impact of MS on walking. The items were generated from 30 patient interviews, expert opinion and literature review (Hobart *et al* 2003). The original psychometric evaluation was based upon data generated by 602 people with MS, and was assessed for data quality, scaling assumptions, acceptability, reliability and validity. Each section is rated from one to five with Likert type responses and has a recall period of two weeks (Bethoux and Bennett 2011). The MSWS-12 is easy to use, inexpensive and takes a few minutes to complete.

Whilst concerns have been expressed by some authors about using patient reported rating scales due to their subjective and potentially biased nature (Guralnik *et al* 1989), the careful development of instruments utilising modern test theory can reduce the relevance of such concerns (Myers *et al* 1993). The MSWS-12 is one such instrument. (Hobart *et al* 2003).

Purpose: To capture the patient's perspective by using psychometrically validated methods. The MSWS-12 assesses different aspects of mobility such as walking, running and climbing stairs (Kieseier and Pozzilli 2012) and is considered to reflect what a person may find difficult over the course of a time period; which cannot be captured with clinician rated short walking tests (McGuigan & Hutchinson 2004; Bethoux & Bennett 2011).

ICF domain: Activity

4.4.2 Psychometric properties of the MSWS-12

Reliability

A range of studies demonstrate good performance regarding reliability (see table 13 below).

Author **(date order)**	**MS Sample** **(unless otherwise stated)**	**Reliability** **(CI included if reported in publication)**
Hobart et al (2003)	n=602 community dwelling,	Cronbach's alpha =0.97 ICC=0.94
McGuigan and Hutchinson (2004)	n=149 community, n = 53 hospital outpatients	Cronbach's alpha = 0.97
Holland et al (2006)	n= 120 range of neurological rehabilitation inpatients	Cronbach's alpha = 0.94
Motl and Snook (2008)	n= 133 recruited from support groups	Cronbach's alpha = 0.97
Learmonth et al (2013)	n=82	ICC 0.97 (95% CI=0.88-0.95)
ICC= intra-class correlation co-efficient, CI= confidence intervals		

Table 13: Studies investigating the reliability of the MSWS-12

Validity

The MSWS-12 has been extensively validated. Its validity was primarily established by Hobart et al (2003), developer of this measure. Table 14 summarises a range of other studies which also support its convergent validity. Further research has found that scores are moderately correlated with physiological measurements of gait including: walking speed (r=-0.59), cadence (r=-0.50), step length (r=-0.53), step time (r=-0.46) and percentage time spent in double support (r=0.54) suggesting that the MSWS-12 is associated with spatiotemporal parameters of gait, in addition to walking speed and endurance (Pilutti *et al* 2013).

Reference	MS Sample	Test and Correlation reported with MSWS-12
Hobart et al (2003)	Community dwelling, n=602 EDSS not reported	MSIS-29 (physical): r=0.79 (p value not reported)
	Inpatient, n=78 EDSS not reported	MSIS-29 (physical): r=0.74 Medical Outcomes Study 36-Item Short Form Health Survey (physical functioning): r=0.79 Functional Assessment of Multiple Sclerosis (mobility section): r=0.70 (p values not reported)
McGuigan and Hutchinson (2004)	Community dwelling, n=149 EDSS 0-7	EDSS: r=0.73, MSIS-29: r=0.80 (p values not reported)
	Outpatients, n=53 EDSS 1-7	EDSS: r=0.65 MSIS-29: r=0.87 (p values not reported)
Motl and Snook (2008)	Sample recruited through support groups, n= 133 EDSS 1-8	EDSS: r=0.80 (p<0.01) MSIS-29: r=0.77, (p<0.01) Accelerometry: r=-0.68, (p<0.01).
Gijbels et al (2012)	Inpatient and outpatient multicentre trial, n = 189 EDSS 0–6.5	Six minute walk test: r2 =0.96 , (p<0.01)
Kieseier and Pozzilli (2012)	Literature review	Timed 25ft walk: r=-0.78 (p=0.01)
EDSS =Expanded Disability Status Scale MSIS-29= Multiple Sclerosis Impact Scale		

Table 14: Convergent validity of MSWS-12

Responsiveness

The MSWS-12 has been shown to be more responsive than the EDSS (Hobart *et al* 2003). A review by Bethoux and Bennett (2011) reported that floor and ceiling effects were less than other measures of mobility indicating its suitability for use across a range of disability levels, providing the person is ambulatory. Holland et al (2006) provided evidence that responsiveness of the MSWS-12 was good when used to assess the effects of inpatient rehabilitation as deemed by an effect size of 1.29. Baert et al (2014) reported it to be better at detecting change than the T25fwt, further suggesting that it may be a more appropriate measure than a short walking test in mildly disabled people with MS. In their multi-centre rehabilitation study, they found it to be the most appropriate walking measure for detecting response to physical rehabilitation (Baert *et al* 2014).

Clinically significant change

Differences are reported in the literature as to what magnitude of change is necessary to be deemed clinically significant. Hobart et al (2013) for example suggests a 15% change is required in contrast to Learmonth et al (2013) who suggests a 53% change (equivalent to 22 points) is clinically significant. Given that their sample characteristics were similar, it is likely that these discrepancies exist as a result of the differing methods they used to calculate responsiveness. Hobart *et al* for instance used the smallest SEM as a bench mark for clinically meaningful change, in contrast to Learmonth *et al* (2013) who used the minimal detectable change (MDC) to define clinically significant change.

Scores reported by Baert *et al* (2014) suggested that in a mildly disabled population (EDSS< 4) a clinically significant change was -10.4 when anchored

to the patients perspective of change in walking, and -11.4 points when anchored to the therapist's perspective. For those with moderate disability (EDSS>4 -11.9 (therapists perspective) to -14.1 (patients perspective) points were clinically significant. These figures are broadly in line with those reported by Hobart et al (2003).

4.4.3 Discussion of MSWS-12

The MSWS-12 has been robustly psychometrically tested by various authors in the target population for this clinical trial. The overall consensus is that this is a robust self-reported measure of walking activity limitation. It is worth considering however, when using self-reported measures, non-ambulatory features such as mood and emotional disturbance may influence the self-ratings of ambulatory performance (Hobart *et al* 2003). Using these in combination with physiological measures may better capture a true reflection of walking ability.

4.4.5 Summary of MSWS-12

The robust psychometric properties and short administration time make the MSWS-12 a useful and practical tool for clinical practice and research. It was chosen for use in this clinical trial to capture the participant's perspective of how MS affects their walking ability.

4.5 Measuring Balance

Balance impairments have been identified as a significant problem for people with MS (Frzovic *et al* 2000; Martin *et al* 2006; Cattaneo *et al* 2007; Hebert & Corboy 2013) and may contribute to advancing disability (Hebert & Corboy 2013). Physiotherapy interventions are commonly used to address balance impairments; a recent systematic review reports these interventions to have a small but significant effect in moderately disabled people with MS (Paltamaa *et al* 2012). A multitude of factors can influence balance which include sensorimotor, proprioceptive and vestibular components (Winter 1995). Due to the complex and flexible nature of balance, assessment can be difficult and may require more than one outcome measure in order to capture a true reflection of a person's balance (Tyson & Connell 2009). In this section the outcomes measures used to assess balance in this clinical trial are reviewed; alternative measures (not used) will be discussed briefly.

4.5.1 Functional Reach Test (FRT) both Forward (FFRT) and Lateral (LFRT).

Description: A clinician rated measure of dynamic standing balance (Tyson & Connell 2009). The participant stands adjacent to a wall with shoulder flexed (forwards reach) or abducted (lateral reach) to 90 degrees. The person then leans forward (or laterally) as far as possible without stepping, thus testing the limits of stability. Measurements are taken with a metre rule in centimetres, the therapists first marks the metre rule in the standing position and then again when the participant has reached forward. This is repeated three times and a mean score is used (Duncan *et al* 1990). The dominant arm is recommended to be used in people with bilateral conditions (Tyson & Connell 2009), although

Kage et al (2009) report that two arms can be used to reduce the effect of trunk rotation upon stability during the FFRT. The FRT is considered to be a quick, easy and cheap test to administer in clinical practice and does not require specialist equipment (Tyson & Connell 2009).

Purpose: To measure the limits of standing stability whilst reaching either forwards or laterally, which may be reduced due to impaired postural control.

ICF domain: Activity

4.4.2 Psychometric properties of Forward and Lateral Functional Reach Tests

Reliability

FRT's have been investigated for reliability in differing populations. Initial studies suggest that reliability for the FFRT was high for both test-retest reliability (ICC 0.92) and intra-rater reach measurements (ICC 0.98) in healthy people (Duncan *et al* 1990). Additionally the LFRT was highly reliable with test-retest repeatability (ICC 0.99) in healthy older females (60+ years) (Brauer *et al* 1999). Inter-rater reliability was high for both reach tests (ICC>0.85) in older adults (mean age 80.2 years) (DeWaard and Bentrup 2002). In these studies the 95%CI's were not reported for the ICC values, which may be because this was not common practice in the 1990's. However this limits the interpretation of the mean ICC estimates, which is important when considering test re-test reliability. Consequently the reader must draw upon sample size and number of repetitions performed to evaluate the ICC.

The forward and lateral FRT's have demonstrated to be reliable for use with a range of neurological conditions (Tyson & Connell 2009; systematic review), with just one study specifically investigating this in people with MS (Frzovic *et al* 2000). It is plausible that neurological fatigue could affect the results of the FRT, however high test- retest reliability (ICC 0.89) has been reported for between morning and afternoon measures despite an increase in perceived fatigue (Frzovic *et al* 2000) (n= 14+14 controls). This indicates that time of day and perceived fatigue does not appear to affect its reliability.

The reliability of the FRT has shown to improve by taking the average score over three measures (ICC 0.89-1.00). This was taken into account and a mean of three FRT's was used in this clinical trial. More recently, Lin et al (2012) has also demonstrated improved reliability using a modified ruler with a fixed stop across the hand.

Validity

There has been limited research to investigate the validity of the FRT in people with MS, consequently literature validating the FRT in other populations has been drawn on in this discussion. The first studies were performed over 20 years ago in healthy people by Duncan *et al* (1990) who investigated the validity of both the forward and lateral FRT's, using a metre rule against a force platform measuring centre of pressure excursion, in a sample of 133 healthy people. Validity of the measure was supported by the strong correlation with centre of pressure (COP) excursion (r= 0.71) (p values only reported if documented in the literature). Individual anthropometric measures such as arm length and height were found to be strongly correlated with reach distance (r> 0.80) (Duncan *et al* 1990).

Jonsson *et al* (2003) proposed that the FFRT is a poor measure for the limits of stability. Based on research exploring the relationship between the FRT and whole body kinematics, ground reaction forces and EMG muscle activity, they consider the potential effect trunk rotation may have on reach distance. In a sample of healthy people, a low correlation was reported between reach distance and displacement of COP (r=0.38) and a moderate correlation (r=0.68) between trunk rotation and reach distance (Jonsson *et al* 2003). This was in contrast to Duncan *et al* (1990) and may be explained by differing samples (older in Duncan ***et al's*** study) and differing methodologies. Duncan et al correlated both anterio-posterior (AP) and mediolateral COP whereas Jonsson et al assessed only AP. However, this raises the question as to whether trunk rotation may have a greater effect on reach distance than displacement of the centre of pressure. Kage et al (2009) found, for example, moderately strong significant correlations between a one arm reach and COP excursion ($r=0.60$, $p<0.05$). A one arm reach was used in this clinical trial as per Duncan et al (1990).

The LFRT has been validated against a 3D analysis of hand marker excursion ($r=0.65$, $p<0.05$) and was found to be weakly correlated ($r=0.33$, $p<0.05$) to COP excursion in healthy older females (Brauer *et al* 1999). In addition the FFRT was moderately correlated with both left and right lateral reach ($r=0.65$ and $r=0.52$ $p<0.05$) in older adults (DeWaard and Bentrup 2002).

In a study comparing FRT of people with MS with controls there were significant differences between forward reach distances ($p=0.02$), but not lateral reach distances. It is notable that this was in a population with very mild clinical

disability (EDSS 0-2, median 1.5) suggesting that the FRT is able to identify balance impairments in the absence of marked clinical signs (Martin *et al* 2006).

The FFRT has been investigated against other measures of balance in people with other neurological conditions, demonstrating it to be moderately correlated with the Berg Balance Scale (BBS) ($r= 0.50$, $p<0.05$) and weakly correlated with the Timed Up and Go test (TUG) ($r=-0.20$) in people with Parkinson's Disease (Brusse *et al* 2005). These differences are in line with clinical expectation. They reflect the notion that the BBS and TUG are measuring slightly different constructs to FRT's. The FRT was designed to measure the limits of standing stability (Duncan *et al* 1990) as one component of balance. The TUG is considered to be a measure of mobility which incorporates balance (Paul *et al* 2014) when walking. The BBS involves assessment of many aspects of balance, including single leg standing and tasks with eyes closed. Consequently the FRT may not measure the same aspects of balance as TUG and BBS which would account for the weak correlations.

In Parkinson's disease the FFR test has been shown to have predictive validity for identifying those who fall. A FR distance of < 25.4cm was a predictor of falls risk which had sensitivity of 30% and a specificity of 92% (n=58) (Behrman *et al* 2002) indicating this may not be an ideal test for identifying fallers. In frail elderly people the FRT indicated that a reach distance of 18.5cm had a 75% sensitivity (95% CI 0.46-0.95) and a 67% specificity (95% CI 0.44-0.84) for identifying fallers (Thomas and Lane 2005).

Responsiveness

At the time of writing there was no published literature regarding responsiveness of either the lateral or forward FRT for use in people with MS. Neither was there any literature on the magnitude of change to be expected with Pilates type exercise, although data is available for a range of other neurological conditions (table 15). Drawing upon the pilot research, the mean change scores were 6.4 cm for forward and 6.8cm for lateral reach distance (calculated from published data: Freeman et al 2010), which is beyond the SEM reported by Smithson et al (1998) and Katz-Leurer et al (2009) indicating that the FRT's can detect change post intervention in the target population.

Author	Population	SEM	MDC
Smithson et al (1998)	Parkinson's Disease (no history of falls)	1.56cm	4.32cm
Smithson et al (1998)	Parkinson's Disease (history of falls)	2.91cm	8.07cm
(Katz-Leurer *et al* 2009) Katz-Leurer	Acute Stroke (modified FRT)		Forward modified FRT= 3.7 cm Paretic side modified FRT = 2.3 cm Non-paretic side modified FRT= 2.67 cm
FRT= functional reach test, MDC= Minimal detectable change, SEM= standard error of measurement			

Table 15: Standard error of measurement and minimal detectable change of the Functional Reach Test in populations with neurological conditions.

Clinically significant improvements

The literature review undertaken failed to identify any literature which provided values for either clinically significant improvements or the MCID in any populations .

4.5.3 Discussion of Functional Reach Tests

There is no gold standard measure of balance for people with neurological conditions. The FRT's are considered to be psychometrically robust for use in neurological clinical practice (Tyson & Connell 2009) and have been used in a number of studies to evaluate the effect of exercise interventions which aim to improve balance for people with MS (Kjølhede et al 2012; Paltamaa et al 2012). Preliminary reliability (Frzovic *et al* 2000) and validity (Martin *et al* 2006) in MS has been reported, however the FRT's have yet to be rigorously tested in this patient group. The FRT does however have demonstrable reliability and validity in stroke and Parkinson's disease (Smithson et al 1998; Tyson and Connell 2009), and hence it is not unreasonable to propose that this is also likely to be the case in MS.

Movement strategy may impact upon the validity of the FRT and should be taken into account (Jonsson et al 2003, refer to detailed discussion on page 314). Standardisation of technique to minimise variability and ensure reliability is important and was undertaken in this clinical trial.

There are three other outcome measures which may have potentially been used for assessing balance in this trial; the BBS, the Trunk Impairment Scale (TIS) and posturography. The TIS has demonstrated validity and reliability for

measuring motor impairment of the trunk in people with MS (Potter *et al* 2012). The BBS has similarly been validated for measuring balance impairment in MS (Cattaneo *et al* 2006) as has posturography (Prosperini *et al* 2011). These three outcome measures are discussed in more detail in the discussion chapter page 178. The choice of outcome measures for this trial was intended to be pragmatic and based on the psychometric properties evaluated in pilot research.

4.4.4 Summary of FRT

The FRT's have been employed across a range of populations, including people with MS. They have been found to be reliable, and a number of studies support their validity as measures of balance. To date the responsiveness and MICD in MS has not been established.

4.5 Activities Balance Confidence (ABC) Scale

Description: The ABC scale measures the psychological impact of balance impairment. It is based upon the construct of self-efficacy (Tinetti *et al* 1990). This scale was initially designed as part of a falls efficacy scale in the elderly. It consists of a 15-item questionnaire, each question rating between 1-10. The scores are converted to a percentage with 100% indicating complete balance confidence. The questionnaire takes a few minutes to complete and requires no specialist training (Woodward 2005). Cattaneo et al (2006) reported that the ABC scale was the most psychometrically robust measure of the five balance tests he evaluated in an MS population (see table 16 below).

Purpose: To measure perceived balance confidence from a self-report perspective. It has been suggested that balance confidence is reflective of physical functioning (Nilsagård *et al* 2012) as fear of falling due to reduced balance has been shown to curtail activity (Gunn *et al* 2013).

ICF domain: Participation

4.5.1 Psychometric properties of ABC scale

Reliability

Reliability of the ABC scale in MS has been investigated in two studies. Cattaneo et al (2007) reported high test-retest reliability (n = 25, ICC 0.92, 95% CI =0.80-0.97), and internal consistency was also reported as high by Nilsagård et al (2012) (Cronbach's alpha, α=0.95).

Validity

Originally designed as an outcome measure for use in older people, many of the studies examining its validity have been performed in older adults. Scores were found to be significantly lower for fallers than non-fallers in the elderly ($p<0.01$), supporting its validity as a measure of balance. Furthermore, moderate to strong significant correlations were noted between the ABC scale and BBS ($r=0.80$, $p<0.01$) and reaction time ($r= -0.64$, $p<0.01$) (Lajoie and Gallagher 2004).

Validation was explored by Cattaneo et al (2006) using an Italian translation of the scale in 51 people with MS (table 16). Participants were included if they were able to stand independently and walk six metres with or without an

assistive device, indicating moderate levels of disability. Those with cognitive impairments were excluded. The ABC scale discriminated between fallers and non-fallers better than both the BBS and the Dynamic Gait Index (DGI). Significant differences between fallers and non-fallers were observed on the basis of ABC scale scores ($p<0.01$).

Test	**Spearman's Rho correlation coefficient and significance**
Berg balance scale	$r=0.48$ ($p<0.01$)
Dynamic gait index	$r=0.54$ ($p<0.01$)
Timed up and go	$r=-0.38$ ($p<0.01$)
Hauser de-ambulation index	$r=-0.45$ ($p<0.01$)
Dizziness handicap inventory	$r=-0.70$ ($p<0.01$)

Table 16: Concurrent validity of the ABC (Cattaneo *et al* 2006)

Nilsagård et al (2012) performed a multicentre cross-sectional study of 84 people with MS, defined as mild to moderately disabled (EDSS 1-6), to investigate the validity of the ABC scale. Correlations against six measures of balance and mobility provided evidence to support its validity (table 17), including its ability to discriminate between fallers and non-fallers.

Test	Spearman Rho Correlation Coefficient and significance
Timed up and go	r= -0.61, p<0.001
Timed 25ft walk test	r= -0.63, p<0.001
Four square step test	r= -0.59, p<0.001
Dynamic gait index	r= 0.62, p<0.001
Timed chair stand test	r= -0.61, p<0.001
12 item multiple sclerosis walking scale	r=-0.75, p<0.001

Table 17: Validity of Activities Balance Confidence Scale

(Nilsagård *et al* 2012)

In people with MS, Cattaneo et al (2006) reported a cut-off point (score) of 40% with sensitivity of 65 % and specificity of 77% for discriminating fallers against non-fallers. The most challenging activities were standing on a chair and reaching and stepping on and off an escalator without support. In summary the ABC scale has demonstrated both concurrent and discriminative validity in people with MS.

Responsiveness

Responsiveness has not been reported for the ABC scale in MS. In the pilot research there was a change score of 15.4 points (calculated from published data) (Freeman *et al* 2010) which was not statistically significant, but was greater than the MDC of 13% reported by Steffen & Seney (2008) in Parkinson's Disease. Drawing upon other clinical populations, Friscia et al (2014) reported that the ABC scale is a moderate to highly responsive measure for people who have dizziness.

Clinically significant changes

There is no data published regarding what constitutes a clinically significant improvement in either MS or in any other patient groups. The MCID has not been established. The MDC has been reported to be 13% in Parkinson's disease (Steffen and Seney 2008). More research to establish clinically significant change is needed.

4.5.2 Discussion of ABC

The ABC scale was initially devised to measure balance confidence in the elderly. The scale is easy to use and has been psychometrically tested for use in both research and clinical practice in people with MS, where it has been found to be reliable and valid (Cattaneo et al 2006; Nilsagård et al 2012). However responsiveness of this measure along with clinical significance, SEM, MDC and MCID has not been reported.

The ABC scale is not an objective measure of balance. It measures the confidence of a person with regard to performing tasks which require functional balance. Although perceptions of balance confidence and walking have shown to be closely correlated in MS (Nilsagård *et al* 2012), there may be discrepancies between balance confidence and objective measures of balance. In validation studies (tables 16 and 17), the lowest correlations (r=0.48) were found with BBS and ABC scale which requires the participant to perform tasks which assess balance. There was no literature identified which tested the validity of the ABC scale against physiological measures such as force plate analysis.

Much like the MSWS-12, the ABC scale is subject to the similar cautions expressed regarding patient reported rating scales (Guralnik *et al* 1989). Factors other than balance confidence could influence results, such as mood. The ABC scale may reflect a person's balance confidence over a longer time period, as the person may recall events over, for example the last week, (rather than for example the FRT which is a test of performance rather than capability). However the ABC scale might be influenced by altered perception and memory. An example of this is the item relating to walking on icy sidewalks; it is possible that during winter people may denote less confidence due to easier recall of these situations. Nevertheless, a self-report evaluation of balance confidence is an important outcome to assess and is a useful adjunct in understanding how impaired balance affects people with MS.

4.5.3 Summary of the ABC Scale

The ABC scale has evidence to support its reliability and validity as a measure of balance confidence in MS. Importantly it measures balance from the perspective of the person with MS, which can be supplemented by objective measures of balance.

4.6 Duel task: Visual Analogue Scale (VAS) to measure perceived difficulty in walking whilst carrying a drink

In the pilot research for this clinical trial, experienced clinicians identified walking whilst carrying a drink to be a task which people with MS often find challenging (Freeman *et al* 2010). Using a VAS is one method of measuring perceived difficulty with this task.

Description: The VAS scale is a straight line with end anchors which are labelled with extreme boundaries, such as ‘no difficulty at all’ or ‘unable’. Although used in clinical research since the 1920’s, it began to appear in the literature more commonly since the 1960’s to measure constructs that are both subjective and dynamic in nature such as anxiety, quality of life and pain (Wewers and Lowe 1990). It is both convenient and rapid to administer (Scheffer *et al* 2010). A horizontal, as opposed to vertical VAS has been shown to produce a more uniform distribution of scores (Wewers and Lowe 1990). In this clinical trial we chose a 10 point linear VAS (aka numerical rating scale), where the person circles the appropriate number on the line to reflect how difficult they consider walking whilst carrying a drink.

Purpose: To capture the participant's perceived difficulty in carrying a drink whilst walking, a dual task requiring balance and mobility, in addition to attention and dexterity.

ICF domain: Activity

4.6.1 Psychometric properties of VAS

Reliability

The literature review did not unearth any studies published assessing reliability of the VAS in MS populations. In the absence of this, data from other clinical populations has been drawn upon. A literature review by Wewers and Lowe (1990) described the reliability of the VAS as good, however this was based on literature published prior to 1990 and hence it lacks details such as ICC and 95% confidence intervals. Additionally this was based on literature typically assessing pain, rather than function.

Reliability has been reported for VAS scales as ranging between ICC 0.40-0.80 in people with irritable bowel syndrome (Bengtsson *et al* 2007). Reliability has been reported as high when used to measure satisfaction after hip arthroplasty (ICC 0.95) (Brokelman *et al* 2012); and moderate when used to measure fear of falling in the elderly (n=650), (Scheffer et al 2010). However the authors highlighted that reliability of the VAS may have been higher because participants with cognitive impairments had been excluded. In summary the reliability of the VAS is dependent not only by the clinical population in which it is used, but also by methodology.

Validity

Validity of the VAS in a range of clinical populations has been tested, including MS. In LBP, VAS demonstrated moderate correlations with changes in Roland Morris questionnaire scores (r=0.46) (Harland *et al* 2014). Moderate correlations (r=0.65) using a VAS to measure stress with hospital anxiety and depression questionnaires have been reported (Lesage *et al* 2012). Brokelman et al (2012) reported a strong correlation between a VAS for pain and VAS for satisfaction of hip arthroplasty (r=0.80), however, lower correlations between the Short Form Health -36 questionnaire and VAS for quality of life (r= 0.21) were reported. In summary the validity of the VAS has been researched in differing clinical populations with varying results which depend upon the construct intended to be measured.

In MS a VAS to measure the subjective experience of walking has been validated against previously validated measures of walking. In a sample of 82 ambulatory people, the VAS scores were significantly ($p < 0.001$) and moderately to strongly correlated with EDSS ($r = 0.679$), T25FWT ($r = 0.606$), Six spot step test ($r = 0.729$), two minute timed walk ($r = -0.643$), MSWS-12 ($r = 0.746$), average daily step count using accelerometery ($r = -0.507$) (Filipović-Grčić *et al* 2013). In this one study the VAS demonstrates good concurrent validity for measuring walking in the target population.

Responsiveness

There have been limited studies performed in MS which have assessed the responsiveness of the VAS for dual tasks, and so the pain literature has been drawn upon in this discussion. The VAS has been reported to be more

responsive to assessing pain than other pain assessment questionnaires (Scrimshaw and Maher 2001). Bolton & Wilkinson (1998) reported it to be responsive in assessment of pain with an effect size of 0.77, when patients were asked to report current pain levels. When asked to report their usual pain levels the effect size increased to 1.34 for the VAS (Bolton and Wilkinson 1998).

In the pilot research the VAS for difficulty carrying a drink was identified as one of the measures which most consistently detected change following eight weeks core stability training. Five of the eight people improved on the VAS after the intervention, with a mean change score of 1.5 for the group (calculated from the published data, Freeman et al 2010), thus providing some evidence of the ability of this specific VAS question to detect change after exercise intervention in people with MS.

Clinically significant changes

There is currently no data available as to what denotes a clinically significant improvement in VAS for dual tasks in MS. Literature published on pain has shown that using a VAS in mm increments, the MCID for mild pain = 11 mm (95%CI 4 - 18 mm); moderate pain =14 mm (95%CI 10 - 18 mm) and severe pain = 10 mm (95%CI 6 - 14 mm) (Kelly 2001). Lee (2003) reported that a mean reduction in VAS of 30 mm represents a clinically significant difference in pain severity that corresponds to patients' perception of adequate pain control. In accordance with this Forouzanfar et al (2003) reported that a 30 mm pain reduction on the VAS was clinically significant for people with complex regional pain syndrome. Zisapel & Nir (2003) reported that a change of 10 mm in the 100-mm VAS signifies a clinically significant change in patients sleep quality. These results suggest that, at least with regard to pain, a clinically significant

change in the VAS scale may lie between 10-30mm and may be dependent upon the condition being assessed. It is not possible to know whether this might also be the case for other constructs, such as those examined in this clinical trial.

4.6.2 Discussion and Summary of VAS

Pilot research identified that walking whilst carrying a drink was a task that people with MS experience difficultly in performing (Freeman *et al* 2010), and that a VAS could be used to measure this from the perspective of the person with MS. While Likert scales were another possible measurement option, they are a more time consuming method of gathering participants opinions (Laerhoven *et al* 2007). Of note, the majority of psychometric testing of the VAS has been performed to assess pain and it is acknowledged that assumptions about its psychometric properties may not be transferable to measuring dual tasks in people with MS.

Other than that undertaken in the pilot work for this clinical trial, there is no research that has investigated the validity, reliability and responsiveness of the specific VAS question used. However, VAS's are widely used in research, with evidence to support their reliability and validity in other conditions. They provide a quick, cheap and easy to perform, self-report outcome measure which is simple to add to a measurement battery.

4.7 Summary of outcome measures chapter

Outcome measures for use both in research and clinical practice are required to demonstrate robust psychometric properties. The majority of those used in this clinical trial have established validity and reliability in the target population and setting, with (at a minimum) pilot work providing information about their responsiveness. The measures for this pragmatic clinical trial were also chosen to reflect UK NHS clinical practice. The intention was to gather both clinician-rated and self-reported measures of balance and mobility.

Section one, Chapter Five: Methods: Procedure for the clinical trial

5.1 Introduction

This chapter will describe and explain the methods and procedures employed in the implementation of the multicentre clinical trial. This chapter follows the guidelines from the CONSORT statement for reporting of RCT's for describing the methods and statistical analysis (Moher *et al* 2010). The interventions are reported according to the TIDieR checklist, which provides guidance for reporting intervention studies (Hoffmann *et al* 2014).

5.2 Aim and objectives

In brief this study builds upon the pilot work undertaken (Freeman *et al* 2010) by implementing an adequately powered RCT.

The primary aim was to determine the effectiveness of Pilates compared with a placebo (Relaxation).

Secondary aims were to:

- compare the effectiveness of Standardised Exercises (SE) with Relaxation, and furthermore to compare Pilates with SE.

- explore underlying mechanisms of change with USI.

5.3 Trial design

The study was a multicentre, assessor blinded, block randomised, placebo controlled trial, performed across seven geographically separate locations:

North Lanarkshire (NHS Trust, Glasgow); National Hospital for Neurology and Neurosurgery (University College London {UCL} Hospitals Trust, London); Newton Abbot Hospital (Torbay and Southern Devon Health and Care NHS Trust); The Merlin Centre (Cornwall); Tavistock Hospital (Torbay and Southern Devon Health and Care NHS Trust); and the School of Health Professions, Plymouth University. The clinical trial was initially designed to be performed at five centres but due to unforeseen circumstances, which involved maternity leave, sick leave and termination of staff contracts, two new recruiting centres were set up which were Tavistock Hospital and the Merlin Centre (a charitably funded MS centre in Cornwall).

The trial was registered on 5th August 2011 with ClinicalTrials.gov (https://clinicaltrials.gov/ct2/show/NCT01414725), trial registration number: NCT01414725. Ethical Approval was granted from the National Research Ethics Service, South West 3 Regional Ethics Committee (REC Reference Number: 10/H0106/88), and from the Faculty of Health Ethics Committee at Plymouth University (REC Reference Code: MS/ab). National Health Service (NHS) Research and Development approval was given from the participating NHS Centres. Recruitment commenced on 1st September 2011 and ceased on 5th March 2013 when the target of 100 participants was reached. The recruiting period was extended by six months from one year to 18 months due to the aforementioned unforeseen circumstances.

5.4 Participants

5.4.1 Inclusion and exclusion criteria

Inclusion criteria:

Eligible participants were all adults aged 18 or over, with a confirmed diagnosis of MS according to McDonalds Criteria (Polman *et al* 2011).

The Expanded Disability Status Scale (EDSS) (Kurtzke 1983) is used in the vast majority of MS clinical trials. While recognised as having significant limitations as an outcome measure (Hobart *et al* 2000), it nevertheless provides a useful descriptor of overall disease severity, and is commonly used to categorise people in terms of their level of function, primarily based on walking ability.

Setting the inclusion criteria at EDSS 4.0 – 6.5 ensured the sample reflected those ambulant individuals for whom the intervention is typically used in physiotherapy clinical practice. It is unusual for people with an EDSS of < 4 to be referred to a physiotherapist since their mobility function is only minimally affected, and those scoring >6.5 are severely limited in their walking.

Exclusion criteria:

People whose cognitive difficulties could interfere with either the informed consent process or the ability to fully engage in an exercise programme which requires bodily awareness were excluded. This was determined by the Abbreviated Mental Test (Sarasqueta *et al* 2001), where scores ≤6 indicated ineligibility.

Potential participants were excluded if they presented with any medical condition contra-indicating participation in core stability exercises. Those who currently or recently (within past 6 months) participated in core stability exercises or had current involvement in another interventional research study were also excluded. This was based on the rationale that Pilates may influence neuromuscular adaptations (Bird *et al* 2012); hence including people who had been exposed to Pilates may have impacted upon the detection of any effect. Participants were questioned about involvement in exercise in a manner which ensured they remained blinded to the exercise groups they would be randomised to. (e.g. 'have you been to any exercise classes, yoga, tai chi, Pilates, swimming?').

Any participant who suffered a relapse during the course of the trial was withdrawn from the trial automatically to avoid confounding of outcomes due to acute neurological changes and medication. Both the researcher (EF) and Dr Freeman (Clinical and Academic Supervisor) were informed immediately and details of relapse and changes in medication were documented by the centre therapist.

5.5 Recruitment procedure

Participants were recruited to the trial either through the physiotherapy department of one of the participating centres, via an advertisement in the SWIMS (South West Impact of Multiple Sclerosis) research newsletter, or via letter of invitation from the participant's neurologist (see appendix 1 and 2) in the case of those living in the South West. For potential participants who were identified from the physiotherapists existing case load or waiting list, the

approach was made by the centre therapist who provided them with the participant information. Whether the participant consented or declined to participate was documented by the centre therapist, including reasons given for declining.

For those who were made aware of the trial by the SWIMS newsletter, contact details for self-referral were provided in the advert (see appendix x). In addition, the SWIMS co-ordinator (Dr Wendy Ingram) identified people which matched the inclusion criteria from the database. Those deemed as eligible were sent an invitation letter from Professor Zajicek or Professor Hobart (Consultant Neurologists) inviting them to participate in the trial (see appendix 2) by making contact with EF. SWIMS did not pass any personal information of potential participants to the researchers (EF and JF), in accordance with the data protection act (https://www.gov.uk/data-protection/the-data-protection-act accessed 6th January 2015 14:22).

If a potential participant contacted EF expressing an interest in the trial the recruitment process was started. For these potential participants, a brief telephone interview was conducted to ensure eligibility using a telephone questionnaire. If deemed potentially eligible, a participant information sheet was sent (appendix 3) either by post or email. If the inclusion criteria was met an appointment was made for the first blinded assessment and initial US scan (Plymouth centre only). At this stage, the information sheet was presented and the potential participant was given an opportunity to re-read this information and ask any questions. Written consent was undertaken and the potential participant was reminded that they were free to withdraw from the trial at any point. Every therapist taking consent had undertaken Good Clinical Practice

(GCP) research training (http://www.crn.nihr.ac.uk/learning-development/good-clinical-practice/ accessed 6th January 2015 15:25).

5.6 Interventions

The three interventions are detailed below according to the TIDieR guidelines (Hoffmann *et al* 2014). Participants from all groups were seen either in a secondary care physiotherapy outpatient department, a domiciliary setting or within Plymouth University's human movement laboratory/ clinical treatment room. This depended on the recruitment location and convenience to the participant.

Some participants were unable to attend all intervention sessions due to illness/ infections or childcare responsibilities. In these cases, participants were given as many sessions as possible and the reasons for non-attendance were documented. In some cases a two week period was allowed to accommodate for public holidays and therapist leave.

5.6.1 Pilates based core stability training programme (Pilates)

The Pilates intervention consisted of 12 x 30 minute, individualised, face to face, Pilates based exercise sessions, which were designed to be delivered on a weekly basis over a 12 week period.

Pilates exercises were selected from an 'exercise basket' which were formulated to reflect current clinical practice (Freeman *et al* 2010) and can be freely accessed at

http://www.mstrust.org.uk/downloads/core_stability_exercises.pdf. The exercise sessions were carried out by Pilates trained Clinical Specialist Physiotherapists in Lanarkshire, UCL and Newton Abbot and by a Pilates trained Physiotherapist at Plymouth University, the Merlin Centre and Tavistock Hospital. The minimum Pilates training requirement was the level one mat work foundation course by an accredited Pilates training body. All therapists had experience of treating people with neurological conditions. In line with the pragmatic approach of this clinical trial the Pilates exercises were not stringently standardised. Therapists chose the exercises and number of repetitions that they considered appropriate for the individual based on their clinical experience and Pilates training.

Participants were assessed for ability and impairment on the first session and the individually tailored exercise programme was formulated using exercises from the 'exercise basket'. The within session therapist assessment of participant was not standardised or documented for this trial. It was recorded in the patient hospital notes, but as this was not a requirement of the trial it was not included in the data collection.

The number of repetitions of each exercise was prescribed according to individual factors such as exercise tolerance and fatigue. As is standard practice, therapists were permitted to use a 'hands on' approach if deemed necessary; for example, to stretch prior to exercises, for postural correction and to promote body awareness in recruiting the deep abdominal muscles. This was considered important by therapists involved in performing the pilot research and adds ecological validity to the study.

Participants were taught to activate the deep abdominal muscles and encouraged to maintain this activation throughout the exercises. The ability to perform and maintain spinal alignment as taught in the Pilates method (McNeill 2014) was not recorded or standardised. It was deemed that the Pilates training of the therapist was sufficient to assess and teach this aspect of the programme. Difficulty of the exercises was progressed over the 12 weeks according to individual response. Progression approach varied, and included increasing repetitions, increasing the difficulty of exercises and/ or prescribing additional exercises from the 'basket'.

Participants were given a home exercise plan and booklet diary with diagrams and instructions demonstrating the exercises. They were asked to undertake approximately 15 minutes of exercises set by the therapist per day and record this in the booklet. Only exercises taught in the face to face session were given as 'home work'. Clear written instructions were given to participants regarding the number of repetitions and how to perform exercises at home, to ensure exercises were being performed safely and effectively.

5.6.2 Standardised exercise programme

The Standardised Exercise (SE) programme followed the same format as the Pilates intervention with respect to number and duration of sessions (12 x 30 minutes), intervention period (12 weeks), and manner in which the exercises were chosen and progressed according to the individual's ability. The nature of the intervention differed in that a SE programme, intended to reflect routine physiotherapy practice, was delivered.

The exercise programme was based upon exercises aimed to improve lower limb strength, trunk and pelvic stability and balance (Barrett *et al* 2009). There was some overlap in the exercises between the Pilates and SE groups, for example 'supine bridges and supine single leg lifts were included in both of the exercise baskets. These exercises automatically activate the deep abdominal muscles (Kavcic et al 2004; Hu et al 2012), and so to differentiate between the two interventions the therapists were asked not to give any instruction to participants allocated to the SE group to use techniques specifically aimed at voluntary activation of the deep abdominal muscles during any of these exercises.

In line with those allocated to the Pilates group, participants were given a home exercise plan and booklet diary with diagrams and instructions demonstrating the exercise. The approach to this home programme mirrored that described for the Pilates group.

5.6.3 Placebo control intervention: Relaxation sessions

The control intervention consisted of three x 30 minute face to face relaxation sessions at four weekly intervals over the 12 week period. A standardised relaxation script was read out during the face-to-face sessions. This relaxation intervention used a muscle contract-relax technique, as an attempt to blind participants to the intervention group (Dayapoglu and Tan 2012).

In line with the other two exercise interventions, participants were asked to undertake a 15 minute daily home exercise programme. They were given a 15 minute audio CD to listen to daily. Participants were telephoned weekly in an

attempt to control for attention bias as feedback from therapists and service users at the design stage of this trial deemed it unethical to ask participants to attend weekly sessions for this control intervention. The relaxation audio CD was recorded especially for the trial by the MS nurse specialist from the MS Trust (Vicki Matthews). The audio CD consisted of guided relaxation exercise, using a muscle contract-relax technique and visual imagery. There were no features of the technique which were designed specifically for an MS population.

5.6.4 Documentation of attendance and adherence

Therapists were given a sheet to record the sessions attended, exercises performed within each session and number of repetitions. Additionally they were asked to record any changes in medication or reasons for missing sessions.

All participants were provided with a tick box diary and they were requested to record adherence to home exercise sessions. Number of days, exercises and repetitions performed were recorded.

5.7.1 Outcome measures and follow-up

The following standardised, validated outcome measures were taken by a blinded assessor at baseline prior to any intervention (week 0), immediately following the face-to-face intervention (week 12) and one month after the intervention period (week 16) to determine any carry-over effects.

The primary outcome measure was a 10 metre timed walk (10mtw).

The secondary outcomes measured were:-

- Walking velocity (metres per second), calculated on the basis of the 10mtw (Hobart *et al* 2013).
- Functional Reach (forward and lateral), clinician rated measures of balance impairment (Duncan *et al* 1990).
- MS 12 item Walking Scale (MSWS-12), a 12 item self-report questionnaire which measures walking impairment (Hobart *et al* 2003).
- Activities-specific Balance Confidence (ABC) Scale, a self-report questionnaire measuring perceived balance confidence (Cattaneo *et al* 2006b).
- 10 point Numerical Rating Scale to determine the participants' perspective of "Difficulty in carrying a drink when walking", identified as a common problem in people with MS (Freeman *et al* 2010).

The outcome measurement procedure, psychometric properties and rationale for choosing the measures are discussed in full elsewhere (chapter 4 page 102). All measures were collected in a protocolised order:

5.7.2 Ultrasound imaging

Ultrasound (US) scans of the deep abdominal muscles were performed on the first consecutive 22 participants attending the Plymouth University site. The purpose of the scans was to explore the underlying mechanisms of change. USI is used to determine thickness of the muscles at rest and during an automatic activation task and thereby enables an exploration of the impact of the

intervention at the level of impairment. Detailed discussion and description of US data acquisition and analysis is given on page 258 .

In brief, the protocol involved capture of US images of the lateral abdominal wall on three occasions: baseline, 12 weeks and 16 weeks.

5.8 Sample size and power calculation

The sample size calculation was based on detecting a clinically significant difference in the primary outcome measure (10mtw) between the Pilates and Relaxation group. There is general agreement that a 20% change in walking time is clinically meaningful (refer to discussion on page 113) (Kragt *et al* 2006). Using a two-tailed test at the 5% significance level to detect a 20% difference in change scores between the Relaxation (control) and Pilates group, and with a standard deviation of 2.9 seconds change (based on the pilot research data, Freeman et al 2010), 30 participants per group were required to achieve 85% power. The sample size was inflated by 10% to allow for potential withdrawal due to relapse (Pilutti *et al* 2014). In total 100 participants were required to be recruited.

5.9 Randomisation

A computer generated block randomisation procedure was used. The randomisation procedure was performed at Plymouth University by the researcher (EF) for all of the centres. The computer programme generated a randomised sequence totalling 20, evenly distributed between each of the three intervention groups, for each of the five centres. The random allocation

sequence was generated by EF using ‘Random Allocation Software’. http://random-allocation-software.software.informer.com/2.0/.

In total 33 participants were randomised to Pilates, 34 to SE and 33 to Relaxation. Concealment was ensured by using opaque envelopes labelled with the participant centre and number (e.g. Plymouth University 01). Twenty envelopes containing a piece of folded card (to enhance concealment) with the intervention group stated inside were sent to each centre. To optimise the rigor of the blinding of the randomisation process the allocation was confirmed by the centre therapist with the trial co-ordinator.

After the South Tees and Lanarkshire centres interrupted recruiting (due to therapist maternity and sick leave), the envelopes were returned to the researcher to allow for randomisation of the additional participants who were required to be recruited through Plymouth University, Tavistock and Merlin centres. When the therapists at South Tees and Lanarkshire returned to recruiting status the envelopes were replicated and then re-sent to these centres. This resulted in one error in allocation which gave rise to unequal groups (Pilates 33, Standard Exercise 35, Relaxation 32). The error was made by EF at the point of placing the card in the envelope; concealment allocation was therefore retained as this did not impact upon the blinding of the centre therapists.

5.10 Data input, checking and error rate

Data entry was performed by EF who was not blinded to the allocated group. The raw data entry into SPSS was double checked against the data collection

sheets by an independent academic to ensure credibility. Out of the 100 participants data sets, 25 participants data (all three occasions, iei 75 raw data sets) were picked at random to be double checked. This totalled 375 individual data entries. The error rate was 1.06%.

All data relating to group allocation and subsequent intervention was also checked for errors by EF and three separate academic researchers. All LOCF entries were checked by the researcher (EF) and another academic; there were three errors (incorrect entries from raw data) which were corrected. Outliers identified by visual analysis were also checked against raw data sheets by EF.

Finally, all EXCEL summary data was checked for errors by two people (a person unfamiliar with the data and EF). This was no ensure that there were no errors in transferring data from the SPSS output sheets to the summary in EXCEL, to check for decimal point placement and to identify any obvious mistakes.

5.10 Statistical Analyses

The statistical analyses plan was detailed in the protocol (Freeman *et al* 2012) in advance of any data analyses. The data were analysed using IBM SPSS version 20. The primary data analysis was by intention to treat, with full analysis of all participants as randomised. The six participants who relapsed were excluded from the analysis (as specified in the protocol); a further 13 were lost to follow up. To maximise available data, and prior to commencing statistical analyses, the decision was made to impute missing outcome values using the last observation carried forward (LOCF) method (White *et al* 2011). This

approach was chosen based on existing evidence that a significant decline in overall group mobility was unlikely over the relatively short timeframe of this trial (Ytterberg *et al* 2008; Freeman *et al* 2013).

Continuous data was tested for normality using the Kolmogorov Smirnov test to examine whether data satisfied the assumptions for parametric testing. Independent t-tests were performed to compare the mean change scores between groups (e.g. Pilates vs Relaxation); statistical significance was set at $p \leq 0.05$. As detailed in the protocol, adjustments for multiplicity of testing were not utilised, as the primary analysis and primary outcome were clearly defined.

To be confident of the conclusions drawn, two sensitivity analyses were undertaken by removing outliers (visually identified using box and whisker plots), on complete case data, and the LOCF data set. In order to allow for the possible effects of age, years since diagnosis, baseline score, adherence to exercise, and balance and mobility scores an ANCOVA analysis was performed as a secondary analysis with these as covariates.

To provide clinically meaningful data to aid interpretation of the results, the within group effect sizes and percentage changes from baseline were calculated for all outcomes, with effect sizes being interpreted according to Cohen's criteria (Cohen 1988).

5.11 Summary of methods

This chapter provides a detailed account of the methods used in the implementation of this multicentre clinical trial and a summary of the process for obtaining the USI data. CONSORT (Moher *et al* 2010) and TIDieR (Hoffmann *et*

al 2014) guidelines were used as a framework for reporting of methods for the clinical trial and interventions respectively. The plan for data analysis was detailed in the published protocol (Freeman *et al* 2012). The next chapter will report the results.

Section one, Chapter Six: Results of the clinical trial

6.1 Introduction

The results reported here include all of the analyses performed on participant demographics and clinical outcome data. The primary data set used to draw conclusions was obtained using the Last Observation Carried Forward (LOCF) method for handling missing data and the results of this analysis are presented in the main text. Summary results of sensitivity analyses are reported in the text with details in the appendices (as directed in the text). The results are reported with guidance from the CONSORT 2010 statement (Moher et al 2010). Results from USI studies are reported in the appropriate chapters in section two, page 286.

6.2.1 The sample characteristics

One hundred participants were recruited across the seven centres. Table 18 details the sample demographic and diagnostic information. At baseline the groups were demographically similar although there was a higher percentage of females in the Pilates group. The baseline scores were similar for all outcome measures except the MSWS-12, in which the Pilates group baseline measure was higher, indicating less walking ability. Statistical testing to compare baseline scores was not performed as advised by Moher et al (2010). On visual inspection, the baseline scores for the walking measures (10mtw, walking velocity and MSWS-12) indicated that the SE group was less impaired in terms of walking with lower MSWS-12 scores and faster walking speeds (SE group was 3.23 seconds faster, which was greater than 20% which is considered to be clinically significant).

6.2.2 Missing data

Missing data was comprised of; participants lost to follow up (n =1), uncompleted forms (n =2) and questionnaires (n = 6), and inconsistent use of walking aids (n =1). Ten metre timed walk data, in which different walking aids were used at different assessments was not used based on the rationale that walking aid type affects walking speed.

	Pilates (n=33)			Standard Exercises (n=35)			Relaxation (n=32)		
	mean	sd	range	mean	sd	range	mean	sd	range
Age in years	54.0	9.2	31 - 73	54.6	11.5	35 - 77	53.8	9.7	40 - 74
% Female	84% (n = 28)			71% (n=25)			65% (n=21)		
% type of MS									
Relapse Remitting	39.4%			37.2%			37.5%		
Secondary Progressive	36.4%			31.4%			25.0%		
Primary Progressive	24.2%			31.4%			34.4%		
Benign	0%			0%			3.1%		
Years since first symptoms	18.9	11.3	2 - 40	18.5	11.6	3 - 44	20.5	11.0	4 - 45
Years since diagnosis	13.2	10.1	1 - 36	13.9	11.0	0 - 41	12.1	10.7	0.5 – 42
sd= standard deviation									

Table 18: Demographic and diagnostic characteristics of the 100 participants

6.2.3 Falls, walking aids and comorbidities

The number falls in the last three months, walking aids, orthotics, whether functional electrical stimulation (FES) used, and co-morbidities was recorded for all participants. On average participants fell 4.19 (mean) times (sd 13.34,

median = 1, range 0-90, {one participant reported falling every day for three months}).

Twenty one percent of participants used no walking aid, 47% walked with one walking stick, 15% with two walking sticks, five percent with one elbow crutch, one percent with two elbow crutches and 11% with a delta frame. FES was used by four percent and ankle-foot orthotics by 12%.

Comorbidities were reported by 64% of the sample. These included asthma, epilepsy, chronic obstructive pulmonary disease, coeliac disease, diabetes (types 1 and 2), diverticulitis, hypertension, myocardial infarction, migraine, osteoarthritis and rheumatoid arthritis.

6.3 Recruitment, allocation and retention of participants

Figure 6 (page 160) details the recruitment and retention of participants. Recruitment commenced on 1st September 2011. The 100th participant was recruited in August 2013, hence the recruitment period was two years in total. Recruitment extended six months past the original plan of 18 months due to unforeseen circumstances, with one centre therapist taking maternity leave, a separate centre therapist taking sick leave, and a further therapist leaving her post. This impacted on the even distribution of participants throughout centres leading to n=40 being recruited at Plymouth and lower numbers at UCL, Tavistock and Merlin (see table 19). All analysis was by intention to treat with each participant data analysed as randomised.

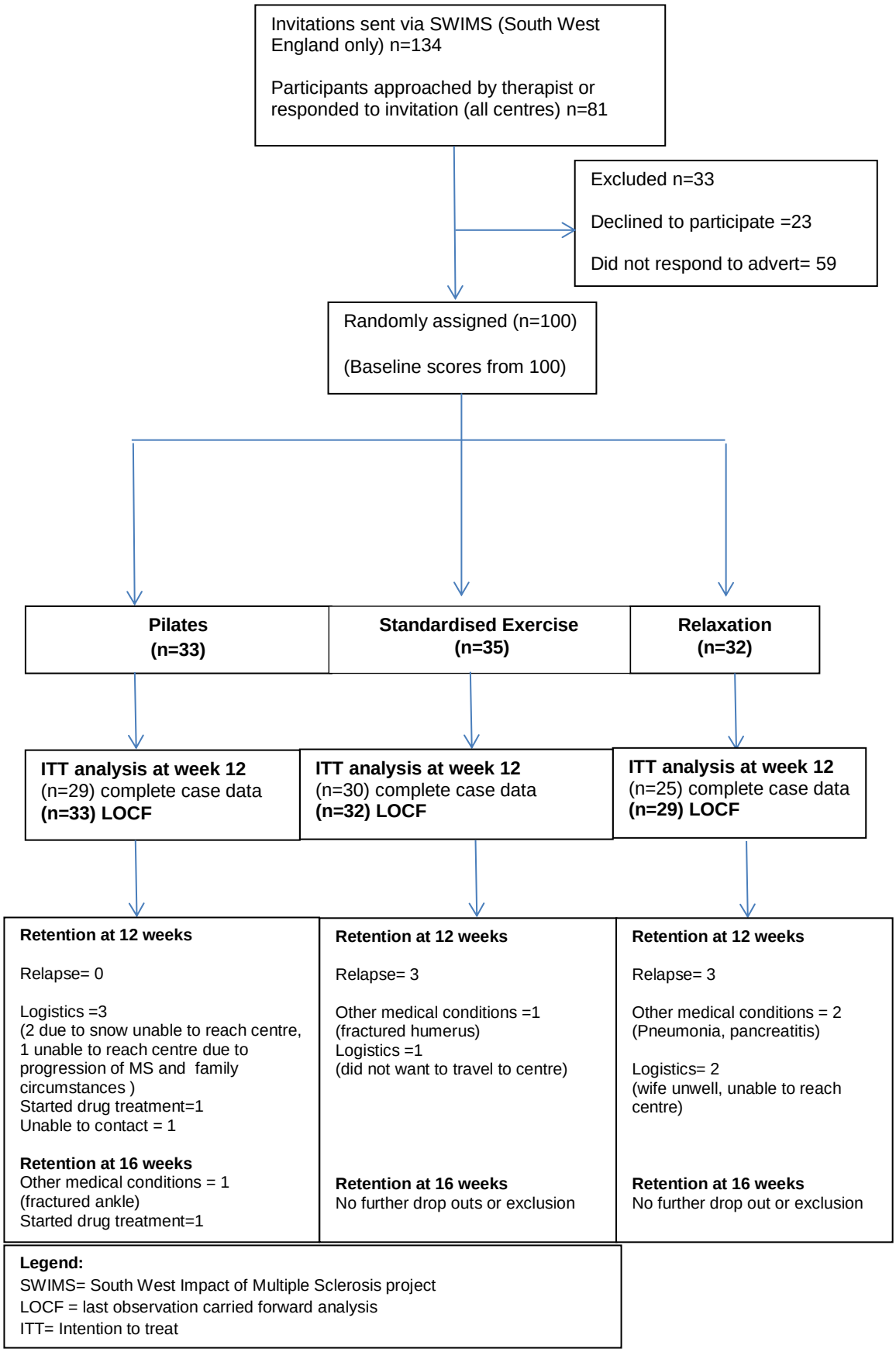


Figure 6 Figure 6: CONSORT flow diagram for recruitment, allocation and retention of participants

Centre	Number of participants
Plymouth	40
Newton Abbot	21
Glasgow, Scotland	19
London UCL	3
South Tees	5
Merlin	7
Tavistock	5
Total	100

Table 19: Distribution of participants in centres

During the trial six participants were withdrawn due to relapse and a further 13 dropped out due to medical and/or logistical problems. There were no reported harms or adverse reactions in any participant that could be attributed to the exercises. The four adverse events were: fractured ankle, fractured humerus (both as a result of falls in the snow, unrelated to the exercise sessions), pneumonia and pancreatitis.

6.4 Within-group changes and between-group comparisons

In order to maintain the statistical power of the sample it was decided (prior to opening the data set) to use the LOCF method for missing data points.

The primary outcome data was tested for normality using the Kolmogorov–Smirnov test and met this and other assumptions required for parametric data testing. The mean within group changes, percentage increases and effect sizes for the primary outcomes are detailed in table 20 (week 12) and table 21 (week 16). Within group change scores denote the difference between baseline scores and follow up scores. The between group differences, p-values and 95% confidence intervals are shown in tables 22 (12 weeks) and 23(16 weeks).

At 12 weeks there were significant differences between Pilates and Relaxation for walking velocity (p=0.04), FFR (p=0.04) and LFR (p=0.04) There were no significant differences between Pilates and Relaxation at 16 weeks. It is worth noting that the sample was only powered to detect changes between Pilates and Relaxation for the primary outcome measure of 10mtw.

At 12 weeks there were significant differences between SE and Relaxation for 10mtw (p=0.05), walking velocity (p<0.01), FFR (p=0.02), LFR (p<0.01), MSWS-12 (p<0.01), and ABC (p<0.01). At 16 weeks significant differences between SE and Relaxation remained for 10mtw (p=0.04), LFR 0.01 MSWS-12 (p=0.03) and ABC (p= 0.03).

There were no significant differences between Pilates and SE at week 12. At week 16 there were only significant differences between Pilates and SE for LFR (p=0.02).

Outcome (week 12)	Pilates		Standard Exercise		Relaxation	
10 metre timed walk	mean	sd	mean	sd	mean	sd
base line score (seconds)	16.16	7.72	12.49	5.05	14.89	6.28
mean change (seconds)	1.72	3.29	2.12	2.23	0.69	3.44
percentage change (%)	9.35	20.21	15.46	13.90	1.38	18.33
effect size	0.22		0.42		0.11	
Velocity						
base line score (m/s)	0.73	0.28	0.91	0.31	0.80	0.35
mean change (m/s)	0.10	0.18	0.17	0.16	0.01	0.13
percentage change (%)	15.90	26.95	21.66	21.35	4.77	20.07
effect size	0.35		0.63		0.03	
Forward functional reach						
base line score (cm)	21.36	10.64	22.20	7.63	20.64	9.27
mean change (cm)	3.10	4.42	4.44	6.97	-0.01	7.18
percentage change (%)	19.99	28.98	26.47	34.75	8.32	25.79
effect size	0.41		0.5		0.1	
Lateral functional reach						
base line score	16.79	5.86	16.11	5.71	16.78	7.25
mean change (cm)	2.15	5.16	3.57	5.23	- 0.84	5.47
percentage change	19.11	44.05	31.15	51.76	1.02	29.88
effect size	0.29		0.57		0.00	
MSWS12 transformed score (0-100)						
base line score	72.15	19.47	58.64	24.45	69.59	20.78
mean change (points)	7.99	16.22	11.67	12.63	2.21	12.39
percentage change	10.26	25.88	21.51	24.79	2.36	20.29
effect size	0.36		0.67		-0.11	
ABC scale (0-100%)						
base line score	3.97	1.54	4.68	2.16	4.27	1.65
mean change (points)	0.66	1.28	1.03	1.27	0.07	1.13
percentage change	17.73	34.17	26.71	33.85	5.52	35.54
effect size	0.43		0.51		0.04	
Walking whilst carrying a drink VAS (0-10)						
base line score	5.53	2.45	5.11	3.05	5.50	2.74
mean change (increments)	0.75	2.36	0.53	1.93	0.17	1.85
percentage change	1.71	83.63	-7.66	120.37	1.28	42.23
effect size	0.31		0.17		0.06	
Note: higher scores = greater ability for velocity, functional reach and ABC scale; lower scores= greater ability for 10mtw, MSWS-12 and VAS scale walking whilst carrying drink						

Table 20: With-in group changes at week 12 assessment

Outcome (week 16)	Pilates		Standard Exercise		Relaxation	
10 metre timed walk	mean	sd	mean	sd	mean	sd
base line score (seconds)	16.16	7.72	12.49	5.05	14.89	6.28
mean change (seconds)	1.72	3.29	2.12	1.53	0.58	3.48
percentage change (%)	7.27	22.88	0.50	42.21	-2.96	30.78
effect size	0.22		0.42		0.09	
Velocity						
base line score (m/s)	0.73	0.28	0.91	0.31	0.80	0.35
mean change (m/s)	0.09	0.21	0.09	0.23	0.01	0.19
percentage change (%)	14.58	30.36	11.31	31.03	4.96	30.08
effect size	0.32		0.32		0.04	
Forward functional reach						
base line score (cm)	21.36	10.64	22.20	7.63	20.64	9.27
mean change (cm)	1.94	6.41	4.09	6.82	1.87	7.14
percentage change (%)	20.13	55.59	27.18	38.40	17.27	33.33
effect size	0.18		0.53		0.19	
Lateral functional reach						
base line score (cm)	16.79	5.86	16.11	5.71	16.78	7.25
mean change (cm)	1.12	5.92	4.70	5.70	0.01	6.46
percentage change (%)	17.97	66.91	42.63	54.89	8.43	38.40
effect size	0.19		0.89		0.00	
MSWS12						
base line score (points)	72.15	19.47	58.64	24.45	69.59	20.78
mean change (points)	3.68	19.72	7.96	15.60	- 0.49	14.26
percentage change (%)	4.12	31.84	16.31	30.41	-3.63	24.28
effect size	0.19		0.35		-0.02	
ABC scale						
base line score (points)	3.97	1.54	4.68	2.16	4.27	1.65
mean change (points)	0.61	1.59	0.74	1.52	0.01	0.99
percentage change (%)	16.75	36.69	19.31	36.85	4.56	31.77
effect size	0.39		0.37		0.01	
Walking whilst carrying a drink VAS						
base line score (0- 10)	5.53	2.45	5.11	3.05	5.50	2.74
mean change (increments)	0.22	2.10	0.14	2.22	- 0.21	1.99
percentage change (%)	1.31	45.87	-5.50	62.45	-16.08	68.50
effect size	0.09		0.05		-0.08	
Note: higher scores = greater ability for velocity, functional reach and ABC scale; lower scores= greater ability for 10mtw, MSWS-12 and VAS scale walking whilst carrying drink						

Table 21: With-in group changes at week 16 assessment

Outcome at Week 12		Pilates	Standard Exercise	Relaxation
10 metre timed walk	group numbers	n=33	n=32	n=29
	mean difference with control (seconds)	1.03	1.43	
	p value	0.23	**0.05**	
	confidence intervals lower	-0.68	-0.04	
	confidence intervals upper	2.75	2.9	
Velocity	group numbers	n=33	n= 32	n=29
	mean difference with control (m/s)	0.08	0.16	
	p value	**0.04**	**<0.01**	
	confidence intervals lower	0.00	0.08	
	Confidence intervals upper	0.16	0.23	
Forward functional reach	group numbers	n=33	n=31	n=28
	mean difference with control (cm)	3.11	4.45	
	p value	**0.04**	**0.02**	
	confidence intervals lower	0.11	0.76	
	confidence intervals upper	6.12	8.15	
Lateral functional reach	group numbers	n=32	n=31	n=27
	mean difference with control (cm)	2.98	4.4	
	p value	**0.04**	**<0.01**	
	confidence intervals lower	0.21	1.59	
	confidence intervals upper	5.76	7.22	
MSWS-12	group numbers	n=31	n=31	n=29
	mean difference with control (points)	5.77	9.46	
	p value	0.13	**<0.01**	
	confidence intervals lower	-1.73	2.99	
	confidence intervals upper	13.27	15.93	
ABC scale	group numbers	n=32	n=31	n=29
	mean difference with control (points)	0.59	0.96	
	p value	0.06	<0.01	
	confidence intervals lower	-0.03	0.34	
	confidence intervals upper	1.21	1.58	
Walking whilst carrying a drink	group numbers	n=32	n=32	n=29
	mean difference with control (increments)	0.58	0.36	
	p value	0.29	0.46	
	confidence intervals lower	-0.52	-0.61	
	confidence intervals upper	1.67	1.33	

Table 22: Between group comparisons at week 12 assessment

Outcome at Week 16		Pilates	Standard Exercise	Relaxation
10 metre timed walk	group numbers	n=33	n=32	n=29
	mean difference with control (seconds)	1.14	1.53	
	p value	0.19	0.04	
	confidence intervals lower	-0.58	0.05	
	confidence intervals upper	2.86	3.02	
Velocity	group numbers	n=33	n=32	n=29
	mean difference with control (m/s)	0.07	0.07	
	p value	0.16	0.19	
	confidence intervals lower	-0.03	-0.04	
	Confidence intervals upper	0.18	0.19	
Forward functional reach	group numbers	n=33	n=31	n=28
	mean difference with control (cm)	0.07	2.22	
	p value	0.97	0.22	
	confidence intervals lower	-3.40	-1.42	
	confidence intervals upper	3.55	5.86	
Lateral functional reach	group numbers	n=31	n=31	n=27
	mean difference with control (cm)	1.11	4.69	
	p value	0.50	0.01	
	confidence intervals lower	-2.14	1.49	
	confidence intervals upper	4.37	7.89	
MSWS-12	group numbers	n=33	n=32	n=29
	mean difference with control (points)	4.17	8.45	
	p value	0.35	0.03	
	confidence intervals lower	-4.68	0.77	
	confidence intervals upper	13.03	16.14	
ABC scale	group numbers	n=32	n=31	n=29
	mean difference with control (points)	0.59	0.73	
	p value	0.09	0.03	
	confidence intervals lower	-0.09	0.06	
	confidence intervals upper	1.28	1.4	
Walking whilst carrying a drink	group numbers	n=32	n=32	n=28
	mean difference with control (increments)	0.43	0.35	
	p value	0.42	0.52	
	confidence intervals lower	-0.63	-0.74	
	confidence intervals upper	1.49	1.45	

Table 23: Between group comparisons at week 16 assessment

6.5 Ancillary Analyses

Analysis of covariance (ANCOVA) was retrospectively performed to determine whether years since diagnosis or adherence to exercises were significant covariates for any of the outcome measures . Baseline scores were significant covariates ($p<0.01$) for 10mtw, walking velocity, both Functional Reach Tests and perceived difficulty carrying a drink, but not for MSWS-12 ($p=0.34$) nor ABC scores ($p=0.65$) at 12 weeks. ANCOVA was not performed at 16 weeks. Due to uneven distribution of participants between centre ANCOVA was not performed using centre as a covariant.

6.6 Sensitivity Analysis

Sensitivity analysis was performed on three variations of the data set: i) complete case data, ii) complete case data with outliers removed on the basis of the box and whisker plot visual analysis, iii) LOCF with relapses and outliers removed. A summary of the significant differences is reported here, all the change scores, standard deviations, mean differences, and 95% CI's for each data set analysed are reported in appendix 4, tables 1-6, page 330.

6.6.1 Sensitivity analysis: Differences with principle data set

- **LOCF outliers removed**

  Reported here: significant results which differ from the principle data set (LOCF) There were not significant differences between Pilates vs Relaxation for any outcome measure at 12 weeks, this differed to the principle data for

velocity, FFR and LFR. At week 16 there was a significant difference between Pilates and Relaxation for the ABC, which was not evident for the principle data set.

For SE vs Relaxation at week 12 there were no significant differences between 10mtw and LFR when outliers were removed in comparison to the principle data set.

- **Complete case data**

Reported here: significant results which differ from the principle data set (LOCF). There were no significant differences between Pilates and Relaxation groups for any outcome measure at week 12 or week 16. This differed to the LOCF analysis as there were significant differences for velocity, FFR and LFR at week 12 for the LOCF data set.

For SE vs Relaxation significant differences for outcome measures were the same with the exception of 10mtw which was significant at week 12 and 16 for LOCF and not for complete case data.

For Pilates vs SE there was no differences between LOCF and complete case data at week 12 and 16.

- **Complete case data with outliers removed**

Reported here: significant results which differ from the principle data set (LOCF). Removing outliers did not change the results produced by the complete case data. There were no significant differences between Pilates and Relaxation groups for any outcome measure at week 12 or week 16. This differed to the LOCF analysis as there were significant differences for

velocity, FFR and LFR at week 12 for the LOCF data set. For SE vs Relaxation significant differences for outcome measures were the same with the exception of 10mtw which was significant at week 12 and 16 for LOCF and not for complete case data.

For Pilates vs SE there was no differences between LOCF and complete case data at week 12 and 16.

6.7 Blinding of assessments

Assessors recorded whether they were blind to the participants' group allocation at each assessment; 84% of the assessments performed were blinded to participant group. Nine percent were completely unblinded i.e. they knew which group the participant had been allocated to. Seven percent were unsure of whether the participant was randomised to the Pilates or SE groups. Whilst rigorous attempts were made to blind participants (page 144), the blinding status of the participants was not recorded.

6.8 Attendance and adherence

Attendance at therapy sessions and adherence to home exercise is detailed in table 24.

Group	Pilates	Standard Exercise	Relaxation
Adherence to therapy sessions	66%	84%	92%
Adherence to home exercises	80%	77%	91%

Table 24: Adherence to sessions and to home exercise programme

6.9 Variability of response of the three groups

Figures 7,8 and 9, page 170, 171 demonstrate the variability of the walking time of the 10mtw at the 12 week assessment (note: not all those that worsened were in the Relaxation group).

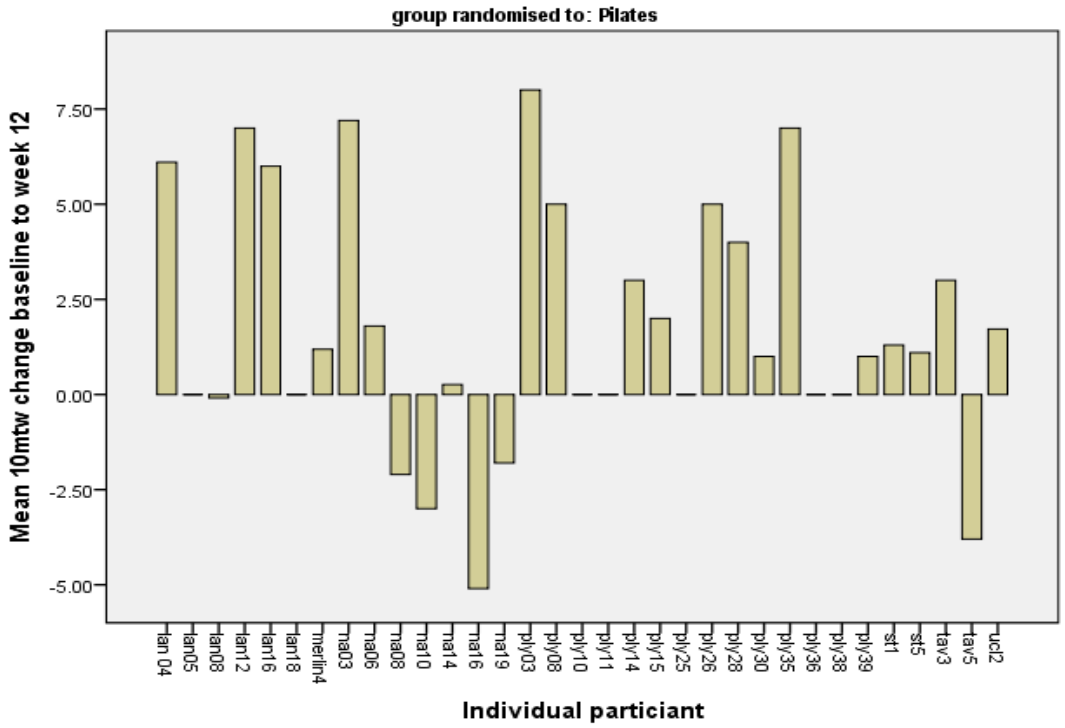


Figure 7: A graph to show variability within the sample for the change in walking speed at the 12 week assessment (Pilates group)

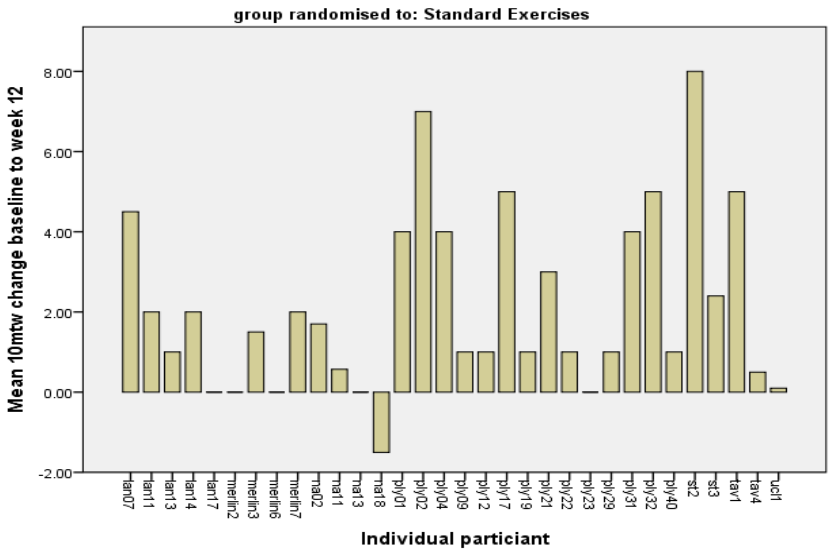


Figure 8: A graph to show variability within the sample for the change in walking speed at the 12 week assessment (Standard Exercise Group)

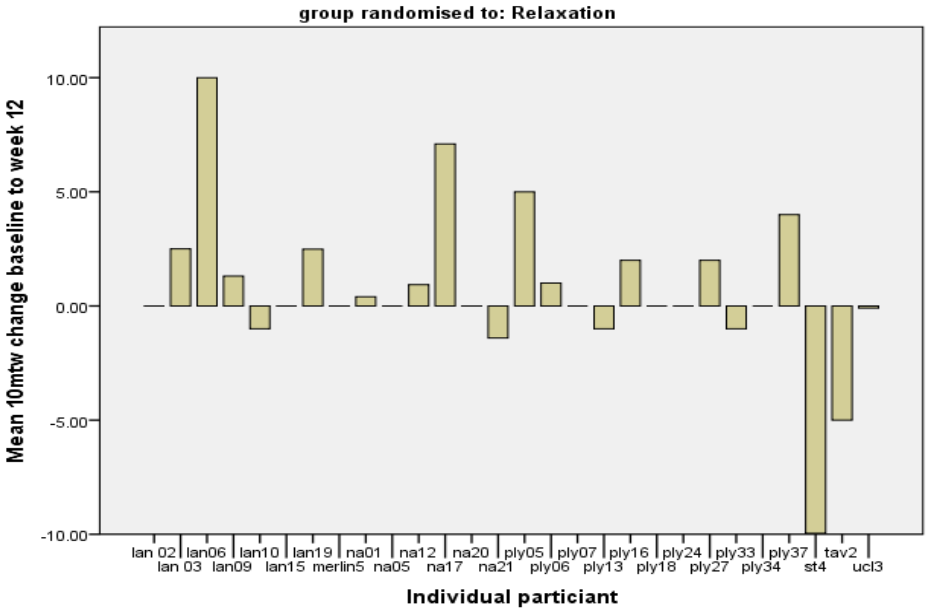


Figure 9: A graph to show variability within the sample for the change in walking speed at the 12 week assessment (relaxation group).

6.10 Summary of results

The sample of 100 participants was powered to use independent t-tests to detect differences between Pilates and Relaxation at week 12 ; while allowing for a 10% relapse rate. The LOCF data set was used for the main analysis. There were significant differences between Pilates and Relaxation groups for walking velocity and Functional Reach Tests (clinician rated measures) which were not sustained at week 16. There were significant differences between SE and Relaxation, with the SE group improving more than Relaxation, for all measures except the VAS scale for perceived walking whilst carrying a drink, and these were sustained at week 16 for 10mtw time in seconds, LFR, MSWS-12 and ABC scale. Sensitivity analysis was performed on complete case data and on both data sets with outliers removed. Sensitivity analysis generally supports the conclusions drawn from the LOCF data set with some explainable differences. Explanations for these findings and clinical relevance is explained and discussed in the following chapter.

Section one, Chapter Seven: Discussion of clinical trial findings

7.1 Summary of findings

This was the first powered, blinded randomised controlled trial conducted to investigate the effects of Pilates in people with MS. As part of the trial the opportunity was taken to recruit participants to compare the effects of a programme of Standardised Exercises with the Relaxation and Pilates interventions; although the trial was not powered for these comparisons.

The trial data was analysed using the LOCF technique to impute data lost to follow up. Significant differences were found ($p \leq 0.05$) between the Pilates and Relaxation (placebo) group at 12 weeks for walking velocity and forward and lateral functional reach. These differences were not retained at the 16 week follow up. Significant differences were found ($p < 0.05$) between SE and Relaxation for all of the outcome measures, except perceived difficulty of walking whilst carrying drink. Significant differences between the SE and Relaxation were retained at 16 weeks for walking time and velocity, lateral reach , self-reported mobility (MSWS-12) and confidence with balance (ABC scale). At 12 weeks within group clinically significant improvements were seen for SE for walking velocity and the MSWS-12. Clinically significant changes for functional reach and ABC have not been established for MS. There were no clinically significant improvements noted in the Relaxation placebo group. Sensitivity analysis was performed on both LOCF and complete case data with outliers removed on both instances and the results confirmed the conclusions.

Following is a discussion of the strengths and limitations of the clinical trial, and explanations for the results. Further, conclusions are drawn from the data. The

focus of this discussion will be on changes at the 12 week assessment as the sample was powered to detect changes over this time period.

7.2 Strengths of the trial

7.2.1Methodology

This is the first multi-centre, assessor blinded, powered, randomised, placebo controlled trial conducted to assess the effects of Pilates upon people with MS. Furthermore, this was the first study to explore the changes at the level of impairment by using USI to visualise and measure the deep abdominal muscles. Other studies to date evaluating Pilates and core stability training in MS have been pilot (Freeman et al 2010) or feasibility studies (van der Linden et al 2013), or studies with methodological limitations such as unblinded assessors (Guclu-Gunduz et al 2013) or with questions regarding their statistical power (Marandi et al 2013).

The trial was conducted at seven geographically dispersed sites which increases the external validity and generalisability of the findings. This suggests that the results were not attributable to a single geographical location or therapist/ assessor. The trial was conducted in a pragmatic setting; the seven centres were comprised of four NHS hospitals, one charity MS centre, one university setting, and the participant's houses (via domiciliary visits) in cases where there were difficulties with travel. The therapists delivering the intervention were all formally trained in Pilates and were experienced in working with people with neurological conditions, and more specifically MS.

7.2.2 Blinding of assessors

Every attempt was made to blind the assessing therapists to intervention allocation. Blinding of assessment is essential to minimise bias towards perceived effects of group allocation; and unblinded trials have been shown to have a propensity to larger treatment effects (Wood et al 2008). Blinding of assessors was recorded at every assessment, with the vast majority (84%) remaining blinded. Reasons detailed for unblinding were; participant telling the assessor, the assessor guessing after being handed the exercise diary (the front cover of the relaxation diary differed to the Pilates and SE diary), and on one occasion the centre assessor was unavailable and there was no option other than the therapist (EF) performing the assessment to remain within the 12 week time scale. The assessors did not ever refer to previous scores to further enhance the rigor.

7.2.3 Randomisation

The randomisation procedure could be considered a strength of the methodology (Schulz & Grimes 2002). A computer generated randomisation sequence was employed and the researcher (EF) prepared sequentially numbered opaque, sealed envelopes which contained the treatment allocation. The participant name was written on the front of the envelope prior to opening. To enhance credibility, the researcher confirmed the participant allocation at each centre with the allocation sequence. It is acknowledged that randomisation by an independent person (such as is undertaken by Clinical Trial Units) would have been preferable, however was not within the scope of this project budget.

A single error occurred in the randomisation procedure. Due to delayed recruitment, additional centres were set up mid-way through the trial, which required a replication of the randomisation sequence. As a result of this there was one duplication to the SE group. The accidental duplication was performed by the researcher (EF) prior to sending envelopes out to new centres and therefore remained concealed. The consequence of this duplication was unequal groups at baseline (SE =35, Pilates =33, Relaxation=32, total 100)., which is highly unlikely to have impacted on the results.

A further strength of the methodology was the implementation of intention to treat analysis with each participant analysed as randomised (White et al 2011). To date, other trials assessing Pilates have not employed such rigor either in randomisation or intention to treat analysis (Guclu-Gunduz et al 2013; Mokhtari et al 2013; van der Linden et al 2013).

7.2.4 Clinical relevance

Over the last decade Pilates and core stability training have grown in popularity within the discipline of neurological physiotherapy (Shea & Moriello 2013). Core stability training is frequently employed as a method of stabilising the trunk in order to improve balance and function in people with MS (van der Linden et al 2013). The original research question was developed in response to a national call from the Therapists in Multiple Sclerosis (TiMS) group to determine areas of interest from therapists working clinical practice. This underlines the clinical relevance of this trial.

7.3 Limitations of the clinical trial

7.3.1 Blinding of participants

In research that evaluates different types of exercise, one of the difficulties that can arise is blinding of participant to treatment allocation. In this trial considerable efforts were made to blind participants to group allocation by using a relaxation-placebo. The fact that the Relaxation group performed a series of progressive muscle contractions allowed the placebo to be described as 'an exercise' in the participant information. However, not all participants were blinded. It was disclosed to the researcher (EF) by one of the centre therapists, that all of the participants were told that they may be allocated to the Relaxation-placebo. This was due to a misunderstanding of the use of informed consent in research. As a result it is impossible to confidently report whether the participants were blinded to group at this centre (n=20 of the total sample). Furthermore, three participants revealed to the researcher (EF) they had guessed that they were allocated to a control group. Conversely, the researcher (EF) noted that the Relaxation exercise served as a good placebo when participants reported that they were 'delighted' to be assigned to an exercise intervention that reminded them of yoga'. Another participant reported that 'this was the most exercise I have done in years'. In retrospect it would have been beneficial to record the blinding of the participants to gain a quantitative evaluation of the success of this blinding process.

7.3.2 Ultrasound protocol

Discussion regarding the USI protocol and the association of the magnitude of abdominal muscle contraction with the Functional Reach Test is detailed on

page 307. To summarise, the USI imaging was performed as an exploratory measure to gain information regarding the underlying mechanisms of change at the level of impairment. This is reported in detail in section two of the thesis.

7.3.3 Choice of outcome measures

The outcome measures were chosen based upon pilot research (Freeman et al 2010). The trial was pragmatic, aiming to replicate UK clinical practice at the time of design. Outcome measures were chosen based both on their psychometric properties (reliability, validity and responsiveness in the target population), and their feasibility for use in a multi-centre trial that was based within a predominately NHS environment (see chapter four page 102 for a full discussion of all outcome measures used). The possibility of different results arising from the use of different outcome measures cannot be excluded. Some potentially alternative outcome measures are discussed below.

Alternative methods of assessing mobility:

Accelerometery

Accelerometery is the use of computer based technology worn by the participant to capture broader activity over a time period. It is considered to be the gold standard for capturing community walking performance as it is performed in a proper ecological setting (Gijbels et al 2010). Using accelerometer based technology it is possible aspects of community ambulation which may not be captured by single occasional tests performed in the clinic. Gijbels et al (2010) used accelerometers over a seven day period to record walking in people with MS and reported that factors such as motivation and

fatigue were accounted for using accelerometry. Accelerometry could have been used as an outcome measure in this clinical trial, however the disadvantages are that it is relatively expensive, is typically not used by clinicians to monitor walking in the NHS, and requires high levels of adherence by participants to ensure accurate results. The use of accelerometry was therefore not in line with the pragmatism of this trial.

Longer walking tests

The 10mtw has consistently been reported to be a valid and reliable measure of walking in people with MS (Tyson & Connell 2009b; Kieseier & Pozzilli 2012). However, since designing this trial research has been published which suggests that longer walking tests may be better equipped to detect change in walking in moderately disabled people with MS. Gijbels et al (2010) found slightly higher correlations with the two minute walk test and accelerometry (r=0.73) than the T25FWT (r=-0.62), suggesting this longer walk test may better reflect “real life” mobility. Gijbels and Dalgas (2012) suggest that for intervention studies a two minute walk test is most appropriate. Baert et al (2014) also reported that two minute and six minute walking tests may be more responsive to clinically meaningful change after rehabilitation than a short walking test. Considering the latest research published, a limitation of this trial may therefore have been the use of the 10mtw test. The two minute walk test may have been a better measure to optimise clinical relevance and responsiveness, while remaining feasible for use within a clinical setting.

7.3.4 Alternative methods of assessing balance

Trunk Impairment Scale and Berg Balance Scale

Other clinical outcome measures of balance which could have been used are the Trunk Impairment Scale and the Berg Balance Scale; both of which are commonly used in neurologically impaired populations (Verheyden et al 2006; Rasova et al 2012).

The Trunk Impairment Scale, whilst originally designed to assess trunk impairment in stroke (Verheyden et al 2004) has been found to be reliable and valid in MS (Verheyden et al 2006) and has more recently been recommended as an outcome measure for use in MS research (Potter et al 2012). This scale is performed seated and measures motor impairment in the trunk muscles. It is possible that this may have better captured changes made in the Pilates group given that a strong focus of Pilates is to train the deep abdominal muscles. It is notable however that the Trunk Impairment Scale is assessed in sitting, in contrast to the functional reach which is performed in standing. It is therefore suggested that the Functional Reach Tests are more likely to reflect functional stability during standing and mobility, although this has yet to be proven.

The Berg Balance Scale is a 14 item test which was designed to measure balance and functional mobility in older adults. It may be the best known measure of balance in adults used by clinicians (Tyson & Connell 2009). This scale has been found to be reliable and valid for assessing non-vestibular balance impairment in MS (Cattaneo et al 2007). Whilst it has been recommended for use in MS research (Potter et al 2012), it has a notable ceiling effect and low sensitivity for discriminating fallers from non-fallers in MS (Cattaneo et al 2007). For these reason, in addition to the reasonably lengthy

time that it can take to administer (in the region of 20 minutes), it was not chosen as an outcome measure.

Posturography

Force platform measures can detect subtle differences in balance deficit which clinical scales may not, with no ceiling effect (Prosperini et al 2011). Using a more responsive, physiological measure of postural stability such as computer based force platform measures (posturography) may have yielded different results, and provided a more in-depth insight into potential differences in the outcomes of these exercise interventions). However, this sophisticated and expensive equipment was not accessible by all the recruiting centres and is not in line with NHS clinical practice. Additionally, properties such as MICD and smallest real change have not been established for posturography in MS (Prosperini & Pozzilli 2013).

7.3.5 Generalisability of findings

The people in this trial were ambulant. Therefore it cannot be specified whether or not any of the exercise interventions used can improve balance and mobility in a more disabled population. Further research is required to substantiate existing evidence from small feasibility studies of wheelchair dependent people with MS (van der Linden et al 2013).

7.3.6 Statistical Analysis

Choosing appropriate methods of statistical analysis is crucial to the correct interpretation of data (Man-Son-Hing et al 2002). The method of data analysis for this clinical trial was designed to be based on two factors; firstly that the data met the assumptions of parametric testing and secondly, upon the power calculation performed, in which there was a comparison between Pilates and Relaxation at the 12 week time point. The analysis plan was designed in conjunction with a medical statistician, protocolised and published (Freeman et al 2012) in advance. The advantage of such an approach is that it increases the transparency in reporting of results of clinical trials and discourages publication of analyses to produce favourable results. Disadvantages are that advances in methods of data analysis are not accounted for.

This clinical trial has been described as a 'placebo controlled trial', however, retrospectively it may have been more appropriate to be described as a 'three armed trial' which encompasses the three intervention groups (Pilates, SE and Relaxation). Similarly, using an approach for statistical analysis which encompassed the interaction between groups and over three times points in which assessment was under taken, such as a mixed factorial ANOVA could have been a more appropriate method. It could also be argued that in using repeated t- tests, Bonferroni corrections should have been performed to correct for multiplicity of testing, thus reducing the risk of a type one error. The use of statistical models to encompass the interactions between time and group, and Bonferroni testing, was discussed at length with the Medical Statistician employed to assist with the designing of the trial and within the supervisory

team. A pragmatic decision was made to perform all data analyses as detailed in the published protocol (using independent t-tests and not performing Bonferroni corrections).

7.4 Explanation of findings

The results of the trial are consistent with the findings of systematic reviews which suggest that exercise is associated with small yet clinically meaningful effects upon mobility (Snook & Motl 2009; Latimer-Cheung et al 2013). Furthermore, a systematic review by Rietberg et al (2004) reported that there is strong evidence to indicate exercise therapy improves mobility. This section will explore and report the explanations of the trial results.

7.4.1 Adherence: attendance at face to face sessions

Supervision by an experienced health professional has been shown to improve adherence to exercise programmes (Garber et al 2011). This is pertinent when designing and progressing exercise programmes for people with MS who experience fluctuations in symptoms and may lack confidence when exercising due to the fear of exercise exacerbating symptoms (Pilutti et al 2014). Interestingly a meta-analysis (of healthy populations) by Rhodes et al (2009) suggested that factors related to exercise prescription (such as intensity, duration and frequency) had very little influence upon the adherence to exercise. Further to this, the type of exercise (i.e. aerobic or resistance) also had a minimal effect upon adherence. The American College of Sports Medicine (ACSM) (2011) recommend that structured supervised exercise

alongside with home based programmes can improved adherence, thereby increasing levels of exercise (Garber et al 2011). This was the approach employed in this clinical trial.

The trial was pragmatic in design and intended to reflect UK NHS clinical practice (Freeman et al 2010). In light of this, and unlike other studies (Gladwell et al 2006), participants were not excluded for missing exercise sessions for reasons such as ill health (e.g. common colds), holidays, bad weather and transport issues. Reasons for non-attendance of sessions are congruent with those described in the study by Learmonth et al (2011). This may have decreased the effect of training but this represents a realistic and achievable exercise programme which can be replicated in clinical practice. Attendance at face to face sessions and the performance of home exercises was recorded. Attendance at therapy sessions for Pilates, SE and Relaxation was 65.5%, 83.6% and 92% respectively. These are expressed as a percentage of possible therapy sessions available to attend. Out of a possible 12 sessions, the mean number attended was 9.7 for Pilates, (median =10, range: 3-12), and 9.8 (median =10, range 4-12) for SE. The reasons for non-attendance at Pilates sessions appeared coincidental and included non-serious illness, holidays and family commitments; none appeared attributable to the contents of the Pilates exercise programme. In the pilot study the attendance (100%) and adherence to home exercise was higher, however the sample size was smaller (n=8) and the intervention period was shorter (8 weeks). This lower attendance at face-to-face Pilates sessions may have impacted upon the results.

7.4.2 Adherence to home exercises

The adherence to home exercises over the 16 week period for Pilates, SE and Relaxation was 79.1 %, 77.7% and 90.7% respectively. Home based training programmes present a very realistic and pragmatic approach to implementation of exercises, however evaluating adherence relies on accurate and honest reporting by participants regarding the volume/ intensity of exercises performed (Dalgas et al 2008). Using a tick box diary with the exact exercise prescription detailed with diagrams, as was employed on this clinical trial, may have improved accuracy of recording.

The adherence for this clinical trial is comparable with other exercise studies in MS. DeBolt & McCubbin (2004) reported mean adherence of 95% to a three times per week home exercise programme over a two week period; the high adherence may be attributed to the short intervention period. Carter et al (2013) reported adherence of 76% at supervised exercise sessions, with participants performing 75% of the 12 week home exercise sessions, equivalent to the adherence of this clinical trial. Romberg et al (2004) reported 93(±46)% adherence to home based exercise over six months.

The adherence to the performance of Pilates and SE was comparable, with higher adherence to the relaxation CD. One possible explanation is that it was easier to adhere to a programme (relaxation CD) which required less physical effort than performing physical exercises. Participants in this trial reported a range of reasons for not performing the exercises. These included: “feeling too tired after being at work all day”; “out with the family”, or because they had already performed physical activities such as “walking around the shops”, “dancing at a wedding” or “looking after grandchildren”.

Enjoyment of both the Pilates and the SE interventions was expressed by participants (reported verbally in sessions and via written comments in the exercise diaries), which mirrors the findings of qualitative research (van der Linden et al 2013). It is common for clinicians to recommend people to choose exercise activities which they enjoy, based on the belief that people are more likely to adhere to this. However, there is limited research to suggest that enjoyment is the factor most likely to promote adherence. Some authors suggest that group based training gives rise to higher adherence and motivation to exercise than home exercise due to the psychosocial and emotional support gained (Romberg et al 2004; Freeman & Allison 2004). Conversely Cattaneo et al (2007) advocate that individualised programmes may better accommodate the high variability of symptoms in people with MS. It is possible that both approaches could be incorporated into programmes such as group based circuit exercises classes which would combine the social benefits of group exercise with individualised programmes.

Dalgas et al (2008) suggested that exercise which does not increase core temperature may provide a more pleasant experience for people with MS, with resistance training being less likely to have an effect on temperature than endurance training. Exercising above the ventilatory threshold has been found to have the most detrimental effect on exercise adherence (Anton et al 2005). Neither Pilates nor SE would be of sufficient intensity to exercise above the ventilatory threshold. Consideration of these factors in the design of future research trials is essential.

7.5 Reversibility of training effects

The magnitude of improvements (percentage increase) made by the Pilates group were not sustained at 16 weeks and the effect size (ES) for both the SE and Pilates group decreased at this follow up assessment. The ES for relaxation was minimal and further decreased at week 16. There is strong evidence from multiple RCT's that physiological adaptations to training are reversed upon cessation of training programmes in healthy people Maintaining intensity is therefore important (Garber et al 2011). This is in line with recommendations from a recent systematic review of exercise in MS that ongoing performance of an exercise programme must be emphasised for training effects to be maintained (Latimer-Cheung et al 2013).

This reduction in ES's over time may reflect the cessation of face-to-face sessions with the therapist, or reduced adherence to the home programme, or a combination of both. This is in line with the associated pilot study in which participants made no further improvements after the withdrawal of the intervention, with two of the eight participants deteriorating (Freeman et al 2010). However, there were still significant differences between SE and Relaxation for walking speed at 16 weeks ($p=0.04$) indicating that some of the improvements were sustained, albeit to a lesser magnitude. One explanation for this could be that adherence to home exercises is high when people expect that a therapist will be assessing their exercise diary, but this decreases when left alone to exercise. Detailed examination of the exercise diaries revealed that participants performed the prescribed exercises seven days per week for the 12 week intervention period (with weekly face to face sessions), but often failed to complete the diaries during the four week follow up period. This could explain

why no further improvements were made over this time period. This information is relevant both for clinical practice, and when designing future studies which employ home exercise programmes.

In summary the self-reported adherence to all three groups was high, which is congruent with published research.

7.5 Variability of response

There was great variability of response seen in the sample; some participants improved up to 50% in walking speed from baseline, while others deteriorated over the intervention period (see figure 8,9 and 10 page 186-187, note: not all those that worsened were in the Relaxation group). This impacted on the magnitude of the group mean change; improvements being small but clinically significant. This finding is in line with studies of healthy people; a summary of RCT's suggested that there is considerable variability in an individual's response to a standard dose of exercise (Garber *et al* 2011).

Factors reported to affect variability in response (in healthy people) include: environmental conditions, individual factors, habitual physical activity, fitness level, physiological and genetic variability, social and psychological factors (Garber et al 2011). It is likely that in MS, in addition to these factors, neuronal damage and deconditioning may further impact upon response to exercise.

Heterogeneity within the clinical presentation and course of MS complicates the design and implementation of research into the effects of exercise. Assembling homogenous and adequately powered samples of people with MS is challenging (Karpatkin 2005), as the degree of variability is relevant for sample

size power calculations (Nilsagard et al 2007). Variability in response to interventions in this trial may have been related to the underlying pathology and associated impairments, which may influence capacity for improvement. For instance, Dalgas and colleagues, in their review of MS exercise trials (which comprised people with mild to moderate disability), highlighted that people with lower EDSS scores had a larger capacity for training adaptation and consequent improvement compared to those with moderate disability (Dalgas et al 2008).

In this trial people with walking speeds slower than 1.2m/s at baseline made the greatest improvements in the 10mtw (i.e. more disabled). Twenty nine percent of slow walkers (slower than 1.2m/s) responded to exercise compared to 12.5% of fast walkers (faster than 1.2m/s), walking speed is discussed more fully on page 115. A potential method to help overcome this would have been to employ stratified randomisation according to EDSS scores and /or baseline walking speed scores and/ or clinical course (Rietberg et al 2004; Kahan & Morris 2012). This method has been previously used, for example, by DeBolt & McCubbin (2004) who undertook stratified randomisation by EDSS level. Paltamaa et al (2008) reported that separating the scores for people who improved and worsened, increases the homogeneity of the data. This may have yielded more definite conclusions regarding the effects of the interventions upon specific groups. Even with relatively narrow inclusion criteria (EDSS scores of 5-6.5) compared to the this trial (EDSS 4-6.5), Learmonth et al (2011) reported that there were wide standard deviations in walking speed at baseline indicating heterogeneity within their sample. From a practical perspective it is noteworthy that narrow inclusion and exclusion criteria inevitably impact upon the speed of recruitment, an important factor to consider when having to recruit

larger samples. Paltamaa et al (2012) suggested that the type of MS does not appear to be a crucial factor in balance, although there is an increasing recognition of the need to investigate interventions separately for relapsing remitting and progressive types of MS (Feinstein *et al* 2015).

The effect of exercise interventions on a more disabled population has been less well researched, with little evidence to evaluate the effects upon people with EDSS>6. This clinical trial included people with EDSS 4-6.5. Until relatively recently the ability of a more disabled population to improve walking with exercise has been questioned, due to the extent of greater neural impairment (Dalgas et al 2012), however the latest evidence from reviews (Swinnen *et al* 2012) and exercise trials (Swinnen *et al* 2012; Briken *et al* 2014; Feinstein and Dalgas 2014) suggests that this may not necessarily be the case. For example, exercise in the form of supported treadmill training resulted in near clinically significant improvements in walking velocity (mean 18%) on the T25FWT in a small sample of people with progressive MS (mean EDSS 6.9) (Pilutti et al 2011). This indicates that more disabled people may have the capacity to improve walking with exercise. In our trial data was not collected for EDSS scores across all the centres, hence it was not possible to perform analysis by EDSS scores. In retrospect collecting this data could have assisted in the analysis and interpretation of results.

In summary, variability of response to exercise is well documented in MS; the variability in the results of this trial is similar to other published research.

7.6 The Interventions

7.6.1 Type of exercise: Pilates

The effect sizes and within group percentage changes were of a lower magnitude for the Pilates group than the SE group programme of lower limb strengthening and balance exercises. The following provides some potential explanations for this.

Pilates as a form of exercise was designed in the early 1900's (see page 66 for full description). The traditional Pilates repertoire was intended to strengthen the entire body and improve flexibility and hence has components of both resistance and flexibility. The ACSM (Garber et al 2011) propose that a resistance exercise programme should comprise of dynamic exercises which result in concentric and eccentric muscle activity and recruit multiple muscles. Exercises should be executed with correct form and breathing technique and include abdominal and spinal muscles (Garber et al 2011). The original Pilates repertoire incorporates all of the suggestions of the ACSM.

Pilates was not designed to be used as a neurological clinical intervention and initially gained popularity after being used to assist strengthening and flexibility of dancers (Siler 2000). It is not surprising therefore that considerable modification and tailoring of the Pilates programme is often necessary to meet individual requirements of the person with MS. The aim of Pilates has been described as 'to improve posture, and improve the mind body connection whilst improving efficiencies of recruitment movement patterning and breathing and centre-ing' (McNeill & Blandford 2013, pg 373). It has a heavy focus of training the core and proximal trunk musculature (Brown 2002; Muscolino & Cipriani 2004; Wells et al 2012).

To optimally target impairments and improve functional capacity in people with MS, exercise may need to focus on more than one area. It may need to be comprised of elements of resistance training (Kjølhede et al 2012), aerobic training (Latimer-Cheung et al 2013), sensory motor balance exercises (Paltamaa et al 2012), and to be task specific (Lord et al 1998). Whilst core stability is one factor which contributes towards balance, training the proximal muscles does not address many of the other key MS impairments. For example, foot drop is a common consequence of MS which impacts upon walking (Barrett et al 2009). Therapists in this trial noted that participants in the Pilates group with foot drop often demonstrated improvements in proximal muscle strength (for example in the ability to perform exercises from the Pilates basket such as planks and bridges) which did not subsequently translate into improved walking due to the presence of foot drop. Unfortunately the numbers of participants experiencing foot drop was not documented (only the use of orthotics and functional electrical stimulation was recorded), hence any conclusions relating to this remains speculative. In clinical practice therapists typically use combined interventions, which may include core stability training, lower limb strengthening and specific balance exercises alongside the use of orthotics, electrical stimulation and medications (Freeman 2008). A consequence of the reductionist approach of many clinical trials is that single interventions are more commonly evaluated than packages of therapy. While this has the advantage of minimising confounders, it has the disadvantage that it may not reflect existing clinical practice (Garrett and Coote 2009). Moreover, combination interventions may be more efficacious (Salhofer-Polanyi *et al* 2013). Future MS research could investigate combined interventions, and packages of rehabilitation aimed to increase balance and mobility.

7.6.2 Comparing Pilates with Standardised Exercise (SE)

There were no statistically significant differences between Pilates and SE groups with the exception of lateral functional reach at week 16 (p=0.04). Whilst it is tempting to perform further analyses to assess the effects of SE, the a priori power calculations would be invalid given that the sample size calculations were based on the decision to compare Pilates with control.

A systematic review of exercise in MS by Rietberg et al (2004) reported that there was no evidence to demonstrate that any one type is better for improving mobility and balance. Recently this notion has been further supported by an MS study comparing 12 weeks of Pilates with aquatic exercise (Marandi et al 2013). Resistance training appears to be an important exercise component for improving the functional capacity of people with MS (Dalgas et al 2008; Latimer-Cheung et al 2013). In contrast to Pilates which focuses on core stability, the SE programme included a number of lower limb exercises. With an established relationship between gait speed and lower limb muscle strength (Jones et al 1999), targeting of these lower limb muscles may provide an explanation as to why there was a greater magnitude of change in walking velocity in the SE group compared to the Pilates group. Future studies are needed to investigate this.

7.6.3 Task specificity

Task specificity or a ‘task orientated’ therapy is based on the specificity of motor learning and skill acquisition, detailing that in order to improve a task it must be

practiced (Straudi et al 2014). Research performed in stroke survivors has shown that the adult human brain is capable of significant adaptations, providing that a sufficient dose of exercise is implemented (Jang et al 2003). Neuro-imaging studies have also demonstrated plasticity in the CNS in progressive MS (Tomassini & Matthews 2012). In light of this, rehabilitation interventions that promote cortical reorganisation by implementing task specific components may be beneficial in MS (Straudi et al 2014).

A study by Lord et al (1998) compared task specific training with a facilitation based approach in people with MS. While both groups improved on the 10mtw, there were no significant differences between groups (p=0.51). Lord et al suggest that one of the reasons for improvements in walking in the task specific training group was that the exercises/ training focused upon this activity. Lord et al further suggested that in an upright position the 'recruitment of synergistic muscle activity, activation of somatosensory receptors' and balance mechanisms were operational in ways which reflect walking. These exercises mirrored many of those used in the SE group; they included stepping up on to a step, squats, and standing balance exercises, amongst others in a standing position. This provides another potential explanation for the larger magnitude and longer lasting duration of change in the SE group compared to the Pilates group, where exercises were mainly performed in supine lying, four point kneeling or prone.

The original research question for this clinical trial focused upon whether training the core muscles improves walking and balance in people with MS. The data suggests that interventions which train the lower limbs in functional positions may be of even greater benefit in improving balance and mobility than

those which focus training on the deep abdominal muscles in a supine and quadruped position. A study specifically powered to address this comparison would be needed to confirm this.

7.6.4 Relaxation placebo control

People who participate in an exercise intervention study frequently demonstrate improvement just by virtue of being involved and assessed (Asano et al 2009). Nilsagård et al (2012) reported that people randomised to a non-exercise (control) intervention expressed disappointment and commented on how they had found new motivation to exercise. Nilsagård et al proposed that this effect may have been emphasised by the study requirement for them to attend follow up assessments. This phenomenon may have occurred in the participants of this trial who were randomised to the relaxation intervention. This underlines the importance of using an effective placebo to control for the effect of therapist attention, and equally important to consider when critiquing research which does not use a placebo control (Mestre et al 2014).

Within group analysis demonstrated small non-clinically significant improvements for participants assigned to the relaxation placebo. At week 12, the mean change in walking velocity increased by 4.8%, forward and lateral reach by 8.2% and 1.0% respectively, MSWS-12 by 8.3% and ABC by 5.5%. Of note, at week 16 the mean improvement on the forward functional reach had increased to 17.7% (Pilates group 20.1%, SE 27.2%).

An MS study by Dayapoglu & Tan (2012) used a nurse led progressive muscle relaxation technique intervention plus a CD for home use (i.e. similar to the relaxation placebo used in this trial) and reported significant within group improvements in sleep quality ($p<0.001$) and fatigue severity scale scores

(p<0.001). While Dayapoglu & Tan (2012) described the muscle relaxation technique as an ‘exercise’, it is highly unlikely that lying supine contracting and relaxing the muscles would be of the necessary load or intensity to generate the physiological changes required to gain sufficient neuro-muscular strength to improve mobility (Latimer-Cheung et al 2013). On the basis of Dayapoglu’s results, it is plausible that fatigue may have improved in the Relaxation group participants which may have enabled them to increase their daily physical activity. Recent research has demonstrated a relationship strong relationship between centrally driven fatigue and balance in MS (r=-0.78) (Hebert & Corboy 2013).This could provide an explanation as to why small (albeit clinically insignificant) improvements were measured in the Relaxation control. Neither fatigue nor sleep quality were measured in this clinical trial and hence no relationship can be determined.

A potential limitation of the trial was that the Relaxation group received only three face to face sessions with the therapist (one per month) compared with the Pilates and SE groups who received weekly (12) sessions. Attempts to match for attention were made by therapists telephoning participants on a weekly basis. However there is a theoretical possibility that the difference in results for the intervention and control groups may have been attributable to differences in therapist attention.

7.6.5 Dose of exercise

Response to exercise interventions is in part determined by the dose of exercise, which is described as the intensity, duration and frequency (Rietberg et al 2004). The prescribed dose for the Pilates and SE group was 12 x 30

minute face to face sessions with 15 minutes of daily home based exercise; and for the Relaxation group was three x 30 minute face to face sessions with 15 minutes of daily home based relaxation listening to the CD. Intensity of exercise can be determined by heart rate or Repetition Maximum (RM) (Collett et al 2011) and was not measured in this trial. The intensity of Pilates training is not well documented in the literature, however, the Pilates and SE interventions used could be reasonably described as low intensity exercise.

Healthy people

There is data to support a dose–response relationship with physical activity and health benefits in healthy people (Garber et al 2011). The exact amount required to generate change is determined by the aims of the individual and baseline levels of physical activity. The number of repetitions, sets and progression dictate the physiological response. To maximise efficacy, training programmes are best tailored to the individual (Mayo *et al* 2013). The ACSM Position Stand document (Garber et al 2011) whose recommendations are based on evidence from RCTs, advocate that for resistance training, an intensity of 40-50% of the one repetition maximum (1RM) is sufficient to improve strength in sedentary healthy people, with eight to ten repetitions adequate to improve strength and power in most adults. This is classified as very light to light intensity. Further to this the recommended number of sets is two to four. However significant gains in muscle strength have shown to be elicited with just one set in deconditioned people. In order to generate physiological change, exercise programmes need to be performed two to three times per week (Garber et al 2011).The optimal methods of progression for healthy people have not yet been determined.

People with MS

The intensity requited to generate change in people with MS is unclear. In our clinical trial the number of repetitions was tailored to the individual. Participants who reported extreme fatigue sometimes performed as few as four repetitions per exercise, whereas less impaired / fatigued participants performed up to 40 repetitions. It has been recommended that using a whole body programme including four to eight exercises placing priority on the lower limbs (Dalgas et al 2008). A systematic review of exercise in MS by Latimer-Cheung et al (2013) provided robust evidence that eight to twenty weeks of supervised resistance training performed two to three times per week at an intensity of 10-12 RM (approx. 70-80% 1RM) increases muscle strength. Latimer-Cheung et al concluded that there is lower level evidence which suggests that training at a frequency of two to three times per week at 60-80% of 1RM can result in significant strength increases.

The required level of intensity or number of repetitions using the approaches commonly implemented by neurological physiotherapists (e.g. core stability training, task specific training) to generate physiological change has not been established.

Current guidelines suggest that there is insufficient evidence available to provide a minimum prescription of physical activity to enhance mobility for people with MS (Latimer-Cheung *et al* 2013). This trial did not measure physiological parameters indicating intensity of exercise (such as heart rate and

oxygen consumption), and hence it is not possible to draw conclusions regarding the intensity of the intervention. However as the data demonstrates that both Pilates and SE interventions resulted in small clinically significant improvements in balance and mobility it is not unreasonable to assume that the intensity was sufficient. It is not known whether increasing the intensity would have resulted in greater improvements; this is a consideration for future research.

7.6.6 Progression

In both the Pilates and SE groups the therapists progressed the exercises according to the individual's response. The two fundamental principles required for optimising fitness are training progression and training volume, both of which are essential for adaptation (Latimer-Cheung et al 2013). Supervised exercise programmes appear to be more effective as modification and progression is facilitated by the professional. The treating therapists employed in this trial were qualified physiotherapists, with the experience and skills to progress the exercise prescription. Evidence from the tick box diaries demonstrates that the participants were self-motivated to progress the frequency and number of repetitions performed during the 12 week period of intervention. The diaries also highlight instances where progression was not possible due to: exacerbation of fatigue, relapse (as protocolised those who relapsed were withdrawn from the trial), and musculoskeletal injury acquired outside of the exercise intervention, but which may have been related to MS (such as trips and falls). These issues are typical of those experienced by people with MS, and are "part and parcel" of incorporating any exercise programme into daily life.

7.7 Comparing the results with other studies

7.7.1 The sample

The demographic and diagnostic characteristic of this sample of 100 participants is representative of both the SWIMS data base (Zajicek et al 2010) and other studies investigating balance and/or mobility difficulties (Paltamaa et al 2008; Baert et al 2014). This supports the generalizability of these results to people with MS who experience mild to moderate disability.

7.7.2 Methodology and outcome measures

In comparing the results with other exercise studies, direct comparison is hindered by limitations in the consistent reporting of methodology. For example, variability in the implementation of outcome measures (such as the use and reporting of use of walking aids and whether self-selected or fastest walking speed is used in the case of the 10twt) has been noted by other authors (Paltamaa et al 2008; Latimer-Cheung et al 2013). Outcome measures used vary between studies, making comparison difficult. The need for an agreed core set of measurements for use in MS clinical trials has been advocated for over a decade (Rietberg et al (2004). Although a range of International Taskforces (Coenen et al 2011; National Institute of Health 2012) and groups (Paul et al 2014) have tried to achieve this, there remains a lack of consensus regarding the best outcome measures to use.

Many of the studies investigating exercise interventions did not use individualised programmes or individual face-to-face sessions with neuro-

therapists, opting for group sessions (Learmonth et al 2011; Tarakci et al 2013; Garrett et al 2013). Whilst group sessions are thought to be cheaper to implement, no health economic analysis has yet confirmed this. The effect of individualised attention cannot be disregarded when comparing results.

7.7.3 Inclusion and exclusion criteria: EDSS scores

The inclusion and exclusion criteria detailed in the methods section (page 143) were defined to select a sample of mild to moderately disabled participants, and was based on pilot research. The EDSS describes and quantifies disability in MS and is very well known and widely used by both clinicians and researchers (Meyer-Moock *et al* 2014). However the EDSS has low reproducibility especially in the lower ranges (Gaspari et al 2002), and poor responsiveness (Hobart *et al* 2000). In this trial the EDSS scale was used for screening purposes to ensure people met the inclusion/ exclusion criteria, rather than as an outcome measure. The telephone version (Bowen *et al* 2001) was used for scoring it at the Plymouth centre as people were recruited from advertisement (i.e. unknown to the therapist working at the centre). In other centres, this was determined by therapist but not formally recorded.

7.8 Comparing the results by outcome measure

To compare the results with those of others, studies were identified from published systematic reviews and meta-analyses assessing exercise in MS (Rietberg et al 2004; Paltamaa et al 2012; Kjølhede et al 2012; Latimer-Cheung et al 2013). Studies were chosen if they investigated similar exercise

interventions (i.e core stability, physical therapy or resistance programmes, not aerobic exercise) and used similar outcome measures.

7.8.1 Measures of walking

The primary outcome measure: 10 metre timed walk

Statistically significant differences were not demonstrated between Pilates and Relaxation (p=0.23) for walking time in seconds at 12 weeks. There were significant differences between SE and Relaxation (p=0.05). The Pilates group had a mean change score of 1.7 seconds (9.4% increase; effect size (ES) 0.2). The SE group had a mean change score of 2.1 seconds (15.5% increase; ES 0.4).

Walking velocity is calculated dividing the distance (10 metres) by the number of seconds taken to walk this distance. Converting time taken to walk 10 metres to velocity produces a more normal distribution of the data than time. The consequences are that results based on speed are less likely to be influenced by skewed distributions thus making them more interpretable using parametric statistics. Velocity results differ from results based on the time taken to walk a set distance (Hobart et al 2013); as a consequence walking velocity (as opposed to time) is frequently used in the presentation of mobility data (Kempen et al 2011). It is noteworthy that the use of velocity data in the statistical analyses generated different results (refer to results chapter, page 161). Walking velocity at week 12 demonstrated significant differences between Pilates and Relaxation (p=0.04) and SE and Relaxation (p <0.01). The Pilates group improved by 15.9% (not clinically significant) with an effect size (ES) of

0.35 and the SE group improved by 21.7% (ES 0.63). These results are consistent with other MS exercise studies with similar samples, which use either the 10mtw or T25FWT as an outcome measure. The main discussion will focus on changes in the 10mtw as this was the primary outcome measure for this trial.

Effect sizes

At the 12 week assessment the ES for walking speed was 0.35 for Pilates and 0.63 for SE (refer to results page 163-164). Effect sizes are commonly used to assess the magnitude and meaning of changes. They are unit-less which allows comparisons to be made across differing time scales for differing outcomes (Asano et al 2009). To apply meaning, an ES of <0.20 is considered trivial, 0.20-0.50 small , 0.50-0.80 moderate and > 0.80 a strong effect (Cohen 1988). It is noteworthy that when comparing effect sizes, multicentre trials show smaller treatment effects than single centre trials (Dechartres et al 2011).

Various factors influence effect sizes (Snook & Motl 2009). The single greatest influence appears to be the length of intervention; exercise programmes of less than three months have an estimated ES of 0.28, whereas interventions exceeding three months show a dramatically reduced ES of 0.09. Snook & Motl suggest that initially bigger improvements are made in the initial training period, with factors such as loss of interest and decreased adherence over a longer time period potentially accounting for the significant loss of effect. Their meta-analyses demonstrated that whether the exercise session was less than 30 minutes or more than 60 minutes, and more or less than three sessions per week, had minimal influence upon effect size. Congruent with Snook & Motl (2009), in this trial larger ES's were noted at the 12 week assessment and were

reduced at one month follow up. One possible explanation is that there may have been less motivation to continue with home exercises after the therapist contact time ceased. This was reflected in the adherence data which showed that participants were less adherent to completing the exercise diary after the 12 week intervention period.

Published research investigating similar exercise upon walking speed

The pilot study sample (Freeman *et al* 2010) walked faster at baseline (10.8 (±2.9) seconds) than participants in this clinical trial (Pilates (16.2 ±7.7), SE (12.5± 5.1), Relaxation (14.9 ±6.3)). This may have impacted on the results as it indicates participants in this trial had a greater level of baseline disability which may have influenced the capacity for improvement.

In a systematic review by Snook & Motl (2009) assessing the effect of exercise upon walking in people with MS, the ES's reported were extremely variable ranging from -0.68 to 0.93. In comparing literature it is noteworthy that the magnitude of ES is directly related to sample size and variability, smaller sample sizes may result in greater ES's (Asano et al 2009). This is reflected in the results of Snook & Motl, where the smallest ES was in the largest sample (n=111) using an intervention of outpatient rehabilitation and the largest ES employed group exercises (Snook and Motl 2009).

Lord et al (1998) reported 10mtw change scores of 6.0 seconds (± 4.7) after 15 sessions (in 5-7 weeks) of task specific training. Their ES of 0.73 (not reported by Lord, but calculated from their data) is considerably greater than the SE group (mean change 2.12 seconds; ES 0.63). Is is possible that this

could be explained, at least in part, by their small sample size (n=10) and lack of assessor blinding.

Romberg et al (2004) assessed the effects of a six month progressive home based exercise programme (n=95; EDSS 1-5.5) using (amongst other outcomes) the T25FWT. They demonstrated significant differences between home-based exercise and control (p=0.04, ES exercise group 0.50; ES control group 0.19). Twenty two percent of the sample demonstrated clinically significant improvements in walking speed (i.e. greater than 20%) with a mean time decrease of 12% (95% CI 16-9%, within group change p<0.01). This exercise programme included strength and aerobic training, and exercises in standing 'for imitation of walking patterns' much like the standing exercises used in the SE group.

The SE intervention replicated exercises used by Barrett et al (2009), which were employed as a home exercise programme. In the study by Barrett et al participants (n=44) were randomised to either the exercise group or functional electrical stimulation, and assessed using 10mtw at baseline, week 12 and week 18. Barrett et al reported significant within group improvements (p<0.01) from baseline to week 18 in the exercise group; with five percent change in walking velocity (ES 0.32) at the 12 week assessment. Differences between the week 12 SE results and Barrett et al's. (2009) could be attributable to a number of factors. Although the samples were broadly similar in terms of demographics and EDSS level, the baseline walking speed of the SE group was slightly faster (0.9m/s) compared to Barrett et al's (0.68m/s), and Barrett et al only included those with secondary progressive MS who demonstrated dropped foot. Additionally the assessors were not blinded to invention, and their exercise

intervention was entirely home based and was only progressed at week 6; both of which may account for the smaller percentage change and ES's observed in their study.

Learmonth et al (2011) performed a leisure centre based group exercise intervention (EDSS scores 5.0-6.5). The 12 week intervention was comprised of two 60 minute sessions per week in which a circuit of balance, strength and aerobic exercises were undertaken. Outcome measures (amongst others) were the T25FTWT and ABC scale. The control group was usual care. Results at 12 weeks were equivalent to those in this trial, with a non-statistically significant ($p>0.05$) but clinically significant increase in walking speed of 24% (ES 0.23).

Tarakci et al (2013) implemented a 12 week (60 minutes, three times weekly) group exercise programme, comprising core stability, lower limb strengthening, balance and coordination exercises in a sample of 99 people with MS (mean EDSS 4). The results demonstrated significant within group improvements in walking time for the 10mtw ($p<0.01$), mean increase 2.7 seconds, sd not reported). Whilst the frequency and duration of contact time was higher than in this trial, the length of the study was equivalent (12 weeks).The mean change scores was equivalent for the 10mtw (Pilates 1.7 seconds (± 3.3); SE 2.1seconds (± 2.2)). The slightly higher change scores of Tarakci et al may be attributable to the higher frequency of sessions compared to this trial. Measurement error is another possible explanation. Without the reporting of standard deviations it is not possible to gauge the sample variability.

In summary the results for the primary outcome measure (10mtw) used in this trial were comparable to those of other MS exercise studies. The small

differences could reasonably be attributed to either differing methodologies or measurement error.

Standard error of measurement and clinically important changes

When evaluating increases in walking time and speed it is important to distinguish genuine clinical change from measurement error. Whilst the 10mtw has confirmed validity and is a highly reliable measure of walking speed (Tyson & Connell 2009a), there is some discrepancy regarding the magnitude of change which reflects genuine clinical change as opposed to measurement error. The sample size calculations for this clinical trial were based on a 20% change in 10mtw which many considered the smallest percentage to detect genuine clinically meaningful change (Schwid *et al* 2002; Kragt *et al* 2006; Hobart *et al* 2013; Learmonth *et al* 2013). Others however suggest different values. For example a 33% increase is suggested by Nilsagard et al (2007), while Vaney et al (1996) considers 28%. This variability may be due to different statistical methods used to calculate the minimal clinically detectable change, of different samples from which these values are drawn. In this clinical trial 27.7 % of people demonstrated improvements in walking speed of greater than 20%, while 14.9% improved by more that 33%.

In terms of velocity (as distinct from time), an increase of > 0.17m/s has been suggested to reflect true clinical change (Morris 2002). In this trial the SE group increased their mean walking velocity 0.17m/s at 12 weeks while the Pilates group increased this by 0.10 m/s; indicating that this may not have been a true reflection of change for the Pilates group. Taking into account the reported

measurement error of 2.6 -3.0 seconds for similar samples (de Groot et al 2006), it could be that neither intervention resulted in changes greater than the measurement error.

It has been suggested that an intervention can be recommended if the entire confidence interval (CI) is greater than the MCID; and that the results of a trial are not negative unless the upper CI is smaller than the predetermined MCID (Man-Son-Hing et al 2002). Based on the pilot research, the predetermined MCID for the 10mtw in this sample was 2.9 seconds and a 20% change was considered to be clinically significant. At week 12 the 95% CI for the mean difference with placebo were -0.7 to 2.7 for Pilates and 0.0 to 3.0 for SE. Based on this, the SE intervention could be recommended for improving walking whereas the Pilates could not.

Clinical relevance of walking velocity in relation to activities and participation

Gait speed is considered such an important predictor of function that it has been described as a 'vital sign'(Bohannon & Williams Andrews 2011). Normal gait speed for healthy people ranges between 1.43 m/s for younger males (<49 years old) to 1.24 m/s for older females (> 60 years) (Bohannon & Williams Andrews 2011). Walking speed has implications for participation in everyday activities. Safely crossing the road relies heavily on unimpaired walking speed; the speed required to use UK pedestrian crossings is 1.2 m/s (Asher et al 2012). Asher et al (2012) defined a walking speed of <1.2m/s as a walking impairment. In this clinical trial at baseline 8.5% had walking speeds of slower than 1.2m/s, indicating moderate disability, of these "slow walkers, 9.1 % of the Pilates and 34.4% of the SE group had improved to the extent that they walked

faster than 1.2m/s at the 12 week assessment. This disparity between the groups in improvement beyond this threshold could, at least in part, be accounted for by the faster baseline speed of the SE group, who walked faster by 0.2m/s, which is equivalent to four seconds on the 10mtw.

7.8.2 Self- report measures of walking: MSWS-12

There were no statistically significant differences between Pilates and Relaxation ($p=0.13$) at 12 weeks. There were significant differences between SE and Relaxation ($p<0.01$) at this time point. The Pilates group had a mean change score of 8.0 points (10.3% increase, ES 0.36). The SE group had a mean change score of 11.7 points (21.5% increase, ES 0.67).

Clinically significant changes in the MSWS-12 vary between 15% (Hobart et al 2013) and 53% (Learmonth et al 2013) (reasons for discrepancy in these published results are discussed in detail in the methods chapter, page 120). Reported values for the SEM also differ: 4.5 points (Hobart et al 2013), 5.66 points (Freeman et al 2013) and 8.0 points (Learmonth et al 2013). In line with this the percentage improvements made in the Pilates group are negligible and not clinically significant. In contrast, using the criteria defined by Hobart et al (2013), the SE group made clinically significant changes which were greater than 8.0 points.

These results are congruent with the 10mtw results in that there were significant differences between SE and Relaxation and the percentage change and effect size were greater for SE than Pilates, thus enhancing the validity of

the findings. These results in combination suggest that Pilates did not result in clinically significant changes in walking in this sample.

Published research investigating similar exercise upon MSWS-12

This section will compare the results of this trial with other published data in relation to the MSWS-12. The papers were identified from a meta-analysis (Snook & Motl 2009) and systematic review (Kjølhede et al 2012) which investigated the effect of exercise training on walking; supplemented by a basic literature search.

Nilsagård et al (2012) reported significant within group changes (p=0.01) in the MSWS-12 after 6-7 weeks of twice weekly, 30 minute therapist supervised sessions of Wii balance exercises (n = 84, EDSS not reported). The Wii exercise programme was comprised of specific balance exercises, strength training and yoga poses. Baseline scores were 50.5 ±25.8 (mean change 5.9 ±11.5, ES 0.51). The baseline score of the Pilates group was 72.1 (mean change 8.0), and SE group baseline was 58.6 (mean change 11.7). The higher level of walking disability in our trial, and/ or longer intervention time (amongst other factors) may have accounted these larger changes.

In a multicentre European study (17 centres, n = 290, EDSS ≤6.5), evaluating mobility change in people receiving between three weeks to three months rehabilitation, Baert et al (2014) reported mean improvements of 7.4 points (±19.7; i.e. 8.6% change) on the MSWS-12. Using anchor based methods of responsiveness they suggest that clinically significant changes were 10.4 points when anchored to patients perspective and 11.4 when anchored to perspective

of the therapist. The SE group had a mean change score of 11.7 points which would constitute a clinically meaningful change according to Baert et al

Straudi et al (2014) assessed the effect of 10 sessions of therapy-led, group based (n = 3), task orientated exercise sessions over two weeks, followed by home exercises for three months (n=24, EDSS 4-5.5). The intervention comprised of exercises in standing, walking, step ups and balance exercises. Significant within group improvement ($p<0.05$) in MSWS-12 scores were reported. At baseline scores were 63.1± 14.0, after the two week intervention: 52.4± 14.1 and at three months following the home exercise programme scores were 65.42 ±16.04, which were worse than baseline. Adherence to the home exercise programme was 58.3%. This study suggests that an intensive period of task specific exercises might result in immediate improvements in self-reported walking on the MSWS-12, but the effects are not long lasting. However without a control group it is not possible to draw definite conclusions. In our trial the improvements in the SE group were greater than this and adherence was higher. Weekly individual face to face sessions with the therapist and high adherence to home exercises may, at least in part, account for this.

In summary, in this clinical trial changes on the MSWS-12 were comparable to those published in other MS exercise studies using similar samples. This adds further evidence to support the conclusions of recent systematic reviews that exercise improves mobility, although the most effective type of exercise remains unclear.

7.8.2 Measures of balance: Forward Functional Reach (FFR) and Lateral Functional Reach (LFR) Test

The mean FFR change for Pilates at week 12 was 3.1 cm (20.0%, ES 0.41). There was a significant difference between Pilates and Relaxation (p=0.04). For the LFR there was a mean change of 2.2 cm (19.1%, ES 0.29); with a significant difference between Pilates and Relaxation (p=0.04).

For the SE group the FFR improved by 4.4cm (26.5%, ES 0.50), which was statistically significant between SE and Relaxation (p=0.02). For the lateral FR there was 3.6cm improvement (31.2%, ES 0.57), which was significantly different between SE and Relaxation (p<0.01).

LFR has not been commonly used as an outcome measure or widely studied in MS, and the literature search did not identify any studies using LFR in MS to compare the results. These differ slightly for the forward and lateral reach; the change was greater by 4.7% for lateral reach than forward reach for the SE group. It is plausible that differing strategies are employed to self stabilise when performing these movements.

This section of the discussion will compare the results of the FFR with that of published data. Drawing on the systematic reviews by Paltamaa et al (2012) and Kjølhede et al (2012), studies which used the FFR and the ABC scale were identified; supplemented by a basic literature search. The FFR distances at baseline were generally shorter in this trial (Pilates: 21.4(±10.6)cm, SE:22.2(±7.6)cm and Relaxation:20.6(±9.3)cm) and the changes of lower magnitude (3.1cm, 4.4cm and 0.0cm respectively) compared to those of the pilot study (mean at baseline 24.5 (± 6.6)cm, change 6.4cm). This was also

the case for the LFR (Pilates:16.8 (5.9±)cm, SE:16.1(±5.7)cm, Relaxation: 16.8(±7.2)cm; mean change 2.1cm, 3.6cm, -0.8cm respectively) compared to the pilot study (baseline 24.9 (±9.6)cm; mean change 6.8 cm). This further supports that this trial's sample was more disabled, which may account for the larger changes in reach distance observed in the pilot study.

The literature search did not identify any published data for the SEM, MDC, MCID, or clinical significance for the reach tests in MS. In Parkinson's disease, the SEM has been calculated to fall between 1.6cm to 2.9cm depending on the level of disability (lower disability = higher SEM) In stroke this has been calculated as 2.5cm, and in vestibular disorders 2.3-2.5cm. The MDC ranged from 4.3 - 9.0cm in Parkinson's disease (all data retrieved from http://www.rehabmeasures.org/Lists/RehabMeasures/PrintView.aspx?ID=950, accessed 17th February 2015, 14.44pm). The literature review failed to unearth any evidence on the MCID for the FFR.

When evaluating the published studies discussed in this chapter a cut off point of 2.9cm was used to determine whether genuine change occurred. This criteria was based on calculations reported from a range of studies (http://www.rehabmeasures.org/Lists/RehabMeasures. Accessed 27/08/14, 17.13pm).

Cakt et al (2010) compared two exercise interventions in which participants were randomised to either: cycling plus balance exercises (n=15), home based lower limb strengthening (n=15) or control (n=15). EDSS scores were not disclosed. The intervention was performed twice per week for two months. Outcome measures (amongst others) included the FFR test, which demonstrated statistically significant within group improvements in the cycling

plus balance group ($p<0.05$, mean change 7.3cm ±2.4), compared to the home exercise or control group where there was no significant changes ($p>0.05$). From this research it would seem that integrating balance exercises into rehabilitation are important for improving FR.

An MS study performed by Broekmans et al (2011) (n=38, EDSS mean 4.3), used a resistance training protocol based on ACSM guidelines for older adults, applying relative workloads to improve muscular strength over two x ten week training periods. Resistance exercise to the leg muscles in a seated position were employed at frequency of 5x60 minute training sessions per fortnight, of 50% 1REP max, increasing the volume and intensity over the time period. FFRT significantly increased in this group compared with control ($p<0.05$, mean change 5.9cm ±1.9cm) after 19 weeks of training, indicating that improving strength, even in a non-functional (seated) position can improve FR distances. Interestingly the change was greater for this intervention than for the SE group which could suggest that leg strength may be more important than task specificity in improving FR, however more research is needed to substantiate this.

Sabapathy et al (2011) performed a randomised pilot study (n =16) in which participants were allocated to either eight weeks of twice weekly supervised endurance or resistance training. The resistance programme included squats and lunges, prone and supine core stability exercises and standing balance exercises. Significant ($p<0.01$) within group changes were reported in both groups, with change scores of 1.4cm for endurance and 5.8cm for the resistance. The resistance programme used exercises which were similar to the SE and Pilates interventions but reported slightly higher change scores (1.7cm

greater than the SE group), which may, at least in part, have been due to the higher baseline scores in Sabapathy et al's. sample compared to this clinical trial.

Vore et al (2011) performed a pilot study in a sample of 13 people with MS, (EDSS scores not disclosed), which consisted of an individualised programme of exercises. These included task specific gait training, resistance and aerobic training in addition to balance exercises with therapist supervision. Amongst others, outcome measures included the FR and ABC (data not reported). Mean changes in distance reached were 2.0cm which was not significant ($p=0.26$). This study was not powered and hence a type two error may exist, however 2.0cm is within the range of measurement error. Baseline FR distances in this sample were low (i.e. more impaired; 12.8cm (±6.44)) in comparison to other published studies which could suggest that people who are more disabled may have less capacity to improve in this outcome.

A meta-analysis of the impact of physiotherapy interventions upon balance found that combined resistance and aerobic training improved functional reach distances compared to control (Paltamaa et al 2012). Of these interventions, a significant effect (ES 0.56, 95%CI 0.02-1.11) was reported when outpatient and home based resistance and aerobic training were employed. These data are comparable with the changes recorded in the SE group for both FFR (ES 0.50) and LFR (ES 0.57).

It has been suggested that specificity of exercise is important to improve balance (Paltamaa et al 2012). The literature indicates that resistance, gait and balance training all improve balance in similar magnitudes as measured by functional reach distances, suggesting that the type of exercise may not play as

great a role as originally anticipated. The data from this clinical trial suggests that training the deep abdominal muscles, as is intended with Pilates, may be less important for improving functional stability in standing than implied by proponents of the Pilates method.

Using the aforementioned MS exercise studies as examples, the baseline reach score appears to influence the change score, with less impaired samples at baseline achieving greater improvements in balance. In light of this, it may be that samples could be stratified at the point of randomisation to ensure groups are well matched on this variable. Identifying who responds best to exercise requires further investigation.

7.8.3 Measures of balance: Activities Balance Confidence Scale

There were no statistically significant differences between Pilates and Relaxation in terms of self-reported confidence in balance as measured by the ABC scale ($p=0.06$, mean within group Pilates change 0.66 points, 17.7% improvement, ES 0.43). There were however significant differences between SE and Relaxation ($p<0.01$, within group SE mean change 1.03 points, 26.7% improvement, ES 0.51). Guidance on the ABC scoring method varies, with some authors reporting transformed data expressed as a percentage while others report raw scores (Nilsagård et al 2012).

The between group comparisons and effect sizes (for this trial) suggest that Pilates is not effective at improving balance confidence, whereas SE is. This again could be related to task specificity (the SE group performed more exercises in standing). We did not measure deep abdominal muscle strength in

the entire sample but it is possible that the gains made in core strength did not translate into improved balanced confidence.

Whilst neither the SEM, MCD nor MCID has been established for the ABC scale for MS, data is available in Parkinson's Disease (SEM = 4.0 points; MCD 11.2-13 points), stroke (SEM = 6.8 points) (http://www.rehabmeasures.org/Lists/RehabMeasures/DispForm. Accessed 29/08/14). The change scores for both Pilates and SE were smaller than the SEM for Parkinson's disease indicating that our results might not be clinically significant. However, as the ES was moderate and there were significant differences between SE and Relaxation, it is likely that the SE intervention genuinely improved balance confidence.

A meta-analysis exploring the effectiveness of physiotherapy interventions upon balance found that there were small but significant effects of motor and sensory exercises upon the ABC scale (ES 0.34, 95% CI 0.01-0.67) (Paltamaa et al 2012). It may be that in order to specifically improve balance, retraining of the sensory systems is also required. In a study by Cattaneo et al (2007) conventional exercises were compared with specific motor and sensory exercises. The conventional exercises were described as 'various therapeutic approaches not directly aimed at improving balance' (Cattaneo et al 2007 page 781). In Cattaneo's sample, baseline ABC scores were 38.5 (± 20.4) – 43.9 (± 21.8). On average, after three weeks (10 sessions) the sensory motor training group improved by 2.32 points, 12.55 for the motor training group and 0.9 points for the conventional exercise group. It is possible that this relatively short intervention time may not have been sufficient to gain the degree of strength changes required to improve balance (in their conventional exercise group),

however it is also recognised that specific balance exercises which employ biofeedback, postural control and exercises directed at improving ankle function may also be required to improve balance and balance confidence (Shumway-Cook and Wollacott 2001).

In a sample of 84 people (EDSS not reported, MS impact scale score 72.1 used to assess disability), Nilsagård et al (2012) reported significant within group changes in balance confidence (p=0.02, mean ABC change 5.0 (± 14.4), ES 0.35) after 6-7 weeks of 30 minute x twice weekly therapist supervised sessions of Wii balance exercises. These differences were not significantly different between the intervention and control (p=0.48). This contrasts with Learmonth et al's. 2011 study where significant differences were found on the ABC scale (p=0.001, 42% improvement, ES 0.94) between control and intervention.

7.8.4 Dual task: Perceived difficulty carrying a drink (Visual Analogue Scale)

At 12 weeks the VAS scale of "difficulty in carrying a drink when walking" demonstrated no statistically significant differences between Pilates and Relaxation (p=0.29, Pilates mean change 0.8 points, 1.7% increase, ES 0.31), nor between SE and Relaxation (p=0.46, SE mean change 2.4 points, 7.7% decrease, ES 0.17).

There is a paucity of published evidence to compare this data with. The pilot study (Freeman et al 2010) demonstrated a 0.9 point VAS change score. There are a number of potential explanations as to why the changes in VAS did not reflect those in either the Functional Reach Tests or ABC for either the Pilates

or SE group. This dual task activity requires multiple components of balance, mobility, attention (cognition), upper limb strength and dexterity and sensory feedback. In this trial none of the exercise interventions focused on improving upper limb function, cognitive attention or sensory retraining. Moreover none focused specifically on practising this dual task.

Because values have not yet been determined as to what defines a clinically significant change for this VAS, then it is not known whether these changes were clinically significant. However the small percentage changes and ES's suggest that neither of the exercise interventions dramatically affected the perceived difficulty of walking whilst carrying a drink. Anecdotally some of the participants reported to the researcher (EF) incidences of 'being able to now walk whilst carrying a cup of tea' and 'no longer needing a napkin underneath to catch spills', while others did not. This is reflected by the wide variability of data; the sd for percentage change in the SE group was ±120.4%.

7.9 Comparing results with disease modifying medications

A meta-analysis by Snook & Motl (2009) reported that the effect of exercise interventions was comparable in magnitude with the effect of disease modifying medications upon the rate of progression of MS, at least in the short term. Recent research tentatively suggests that exercise may be able to slow disease progression (Dalgas & Stenager 2012). It is therefore worth comparing the data with that of MS drug trials such as those investigating the effectiveness of Fampridine in improving walking speed. In a drug trial spanning 14 weeks, improvement in walking speed in Fampridine-treated people was 25·2% (95% CI 21·5 - 28·8%) and 4·7% (1·0 - 8·4%) in the placebo group (Goodman et al

2009). In this study the SE group improved by an equivalent 21.7% over the 12 week period. Future research could be directed at combined interventions, for example where exercise is used in conjunction with drug therapy, to determine whether this further enhances the benefits gained.

7.10 Additional factors which may have affected the results

7.10.1 Unforeseen circumstances

There were a number of unforeseen circumstances which impacted upon the speed of recruitment and the distribution of participants amongst centres. At the London centre ethical approval for the clinical trial was very delayed; the therapist's contract expired after recruiting only three participants. At the South Tees centre the therapist took maternity leave mid-way through the trial. At the Scotland centre the therapist ruptured her anterior cruciate ligament and then subsequently fractured her leg mid-way through the recruiting period. As a result of these circumstances, new centres were initiated in Cornwall (Merlin centre) and in Devon (Tavistock hospital). Recruiting of participants was re-started at South Tees and Scotland when therapists returned from leave. As a consequence the trial recruiting period was extended by six months and there was an uneven distribution of participants amongst the centres. These unforeseen circumstances are an inevitable consequence of performing clinical research within a pragmatic setting, and limited budget. The uneven distribution of participants meant that ANCOVA analysis for effect of centre was not able to be performed.

7.10.2 Relapse and dropout rate

Exercise has proven to be safe and well tolerated by people with MS (Dalgas et al 2009 ; Paltamaa et al 2012; Pilutti et al (2014). Indeed Pilutti's systematic review demonstrated that the rate of relapse was lower for participants randomised to exercise compared to control groups (4.6% exercise, 6.3% control). Additionally these reviews demonstrated incidence of adverse effects during exercise was the same in MS and healthy populations. The relapse rate in this clinical trial (which was accounted for in the sample size calculations at an estimated rate of 10%) was six percent and there were no adverse events which related to the exercise.

7.10.3 Thermosensitivity

Thermosensitivity is a common phenomenon in MS, with 80% of people developing neurological symptoms in response to an increase in core temperature. This has been described as a pseudo exacerbation due to a transient increased blockage of nerve conduction in demyelinated fibres (Guthrie & Nelson 1995). The relevance of thermosensitivity to the results of this clinical trial may be linked to the time span of the trial. Firstly involvement of each participant was over a four month period in which the weather sometimes changed considerably. It was noted by the researcher (EF) that over the hotter summer months participants reported increases of fatigue; on occasions cancelling training sessions or not undertaking home exercise as a direct consequence of heat related fatigue. For these people, this decreased the intensity and frequency at which they were able to exercise. Additionally precooling has been shown to increase the speed of walking (White et al 2000).

Participants in this trial who performed their first walking assessment on a cold day and second assessment on a hot day (or visa versa) may have been affected as a result in terms of their walking speeds. The seasonal effect upon walking measures was avoided by Paltamaa et al (2008) who attempted to control for this by taking measures exactly one year apart. It is expected that the randomised controlled design of this trial will have negated the impact of these seasonal fluctuations on the results.

7.11 Predicting who will respond to physical therapy treatments

Large variability in results has been reported in many MS rehabilitation exercise studies and iIt is generally accepted that heterogeneity of response is typical in MS (DeBolt & McCubbin 2004; Sabapathy et al 2011; Karpatkin 2005; Latimer-Cheung et al 2013). This was also the case in this clinical trial, as demonstrated by the variability of 10mtw change scores as illustrated in figure 8-10 page 186. In order to target rehabilitative exercise to best possible effect, identification of people who respond favourably to exercise is required.

Cattaneo et al (2007) suggested, on the basis of anecdotal reports of therapists, that people responding favourably to exercise interventions can be predicted. These (anecdotal) predictors include a lack of prior experience of rehabilitation programmes, only one sensory impairment, a lack of cerebellar involvement, motivation to engage in treatment, and (less importantly) axial muscle strength and fatigue. These align to some degree with findings from a preliminary study (Langdon and Thompson 1999), which identified cerebellar and cognitive (verbal intelligence) function as being influential in determining physical

rehabilitation outcome. Further studies are required to better understand these potential predictors in order to improve targeting of resources.

7.12 Longer term effects of exercise

Few exercise studies have been performed which assess the longer term effects of exercise interventions. Studies tend to assess outcomes immediately post intervention which makes it unviable to draw conclusions regarding their longer term effects (Latimer-Cheung et al 2013). This clinical trial performed outcome measures at baseline, week 12 (directly post intervention) and at a follow up period one month after contact time with a therapist had ceased (with participants being asked to continue with home exercises). It could be argued that a one month follow up period does not constitute long term. Future research is required to provide further evidence as to the long term effects and adherence of exercise programmes, perhaps for as long as one year post cessation of the intervention.

7.13 Summary of discussion chapter

To summarise there are many factors which may have affected the results of the trial. Amongst others, these could be related to the intervention approach, the dose and adherence. The choice and responsiveness of the outcome measures also inevitably affects the results. Using seated measures of trunk stability may have better captured any changes in trunk stability, which is the focus of Pilates. However the outcomes were specifically chosen to best answer the research question which was originally formulated by practising clinicians,

namely “whether these exercise interventions impacted on balance and mobility in ambulant people with MS”. Other factors may also have affected the trial which were unrelated to the design, such as individuals response to exercise and unforeseen circumstances such as relapse rate and thermosensitivity.

Section One, Chapter Eight. Conclusions

8.1 Summary of findings from the literature review of the effects of Pilates

Theories of core stability were primarily introduced in the early 1990's, these proposed that training of the deep abdominals can improve spinal stabilisation. Research has primarily focused upon the onset of TrAb activation and its role in providing a 'corset' to aid core stabilisation. Recent research disputes this assertion. It is now considered that all of the trunk muscles provide degrees of stabilisation via a complex synergistic neuromuscular coordination and co-contraction of agonist and antagonist paraspinal, deep abdominal and trunk muscles. In the literature there is no unanimity as to how to define core stability, and little consensus about the best methods for measuring core stability.

Pilates appears to positively influence balance in people with neurological conditions. Conclusions drawn from the literature are with reservation due to due to the poor methodological design and reporting of studies. Pilates has not demonstrated superiority over other forms of exercise in improving balance. There have, however, been no reported ill effects or harms, hence Pilates can tentatively be considered a safe form of exercise for people with MS.

8.2 Summary of methods

This is the first adequately powered, multicentre, assessor blinded, randomised, placebo controlled trial performed to evaluate the effects of Pilates upon the balance and mobility of ambulant people with MS. The primary aim of this clinical trial was to evaluate the effects of a 12 week programme of Pilates. The trial was powered to detect changes in the primary outcome measure (the 10 metre timed walk test) at week 12 (directly after the intervention period ceased).

Secondary outcome measures were; walking velocity, forward and lateral functional reach, the self-report MSWS-12 walking scale, the self-report Activites and Balance Confidence scale, and a Visual Analogue Scale to determine the patients perceptions of difficulty walking whilst carrying a drink. Given the time and resources to perform the trial, the opportunity was taken to evaluate the effects of a programme of Standardised Exercises. Intention to treat analysis was performed with each participant analysed as randomised, using the last observation carried forward technique. Independent t- tests were used to compare groups at week 12 and then week 16. Ultrasound imaging of the deep abdominal muscles was performed to explore the effects of these exercises at the level of impairment in a sub-sample of participants (results are reported in section 2,chater 4 page 286).

8.3 Summary of results

One hundred participants were recruited and assessed at baseline. Thirteen of these relapsed and were excluded from the analysis as protocolised. Comparing a 12 week programme of Pilates with Relaxation (placebo control) demonstrated neither statistically nor clinically significant (< 20%) between group differences at the 12 (p=0.23) and 16 (p=0.19) week assessments on the primary outcome, the 10mtw. There were, however, statistically significant improvements in the clinician rated measures; walking velocity (p=0.04), forward (p=0.04) and lateral (p=0.04) functional reach at 12 weeks. These were not sustained at the 16 week follow up assessment. The magnitude of these changes was small as defined by the effect size.

Comparing Standardised Exercises with Relaxation (placebo control) demonstrated statistically significant between group differences for the 10mtw ($p=0.05$), walking velocity ($p<0.01$), forward ($p=0.02$), and lateral ($p<0.01$) functional reach, MSWS-12 ($p<0.01$), and ABC ($p<0.01$). The magnitude of these was moderate and improvements in walking velocity were considered to be clinically significant (> 20%); most were sustained at a lesser magnitude at the 16 week follow up.

Comparing Pilates with Standardised Exercise demonstrated no statistically significant between group differences with the exception of the lateral functional reach at week 16 ($p=0.02$). The trial however was not powered to detect differences between these two interventions. Multiple sensitivity analyses were performed and supported the conclusions drawn.

8.4 Summary: Explanations of findings

The results may have been affected by a number of factors which include the type and dose of exercise, levels of adherence and attendance to the exercise programme. The Standard Exercises may represent a more task orientated approach as many of the exercises were performed in standing. It may be that voluntary activation of the deep abdominal muscles as taught in the Pilates method is not a requisite of improving balance and mobility. Furthermore attendance at Pilates session was lower.

Choosing alternative outcome measures which may have been more responsive to measuring balance and mobility in the target population may have

demonstrated different results, however this would have detracted from the pragmatism of the trial.

8.5 Contributions to knowledge

The results of this clinical trial demonstrate that Pilates has a small effect upon balance and walking in ambulant people with MS. The clinician rated measurements of balance and walking (10mtw and FRT's) were significantly different to Relaxation, although the improvements were not considered clinically significant and were not retained at 16 weeks.

In contrast, significant differences between SE and Relaxation were demonstrated in nearly all outcome measures at 12 weeks, were considered clinically significant and generally retained at 16 weeks. In light of this it could be considered that voluntary activation of the deep abdominal muscles, as purported by the Pilates method, are not required to improve balance and walking in ambulant people with MS.

Section One, Chapter Nine: Future research

9.1 Exercise: targeting those who respond favourably

The results of the clinical trial indicate that whilst the effect of Pilates was small and the effect of SE was moderate, the variability in response between individuals was large. Some people improved greatly in the 10mtw in the Pilates group (the participant who made the greatest improvement in walking speed was assigned to the Pilates intervention), whilst other deteriorated. The variability in the improvements made suggests that people with MS may have differing responses to exercise. Improvements may not be entirely dependent upon the type of exercise intervention. Future research would well be directed towards identifying those who respond to exercise interventions and determining reasons for these responses. Initially, understanding differing responses could be enhanced by using the data set generated by this clinical trial; by further evaluating factors such as baseline scores, attendance at sessions and adherence to home exercise. Investigating factors such as type of MS, relapse rate and years since diagnosis may also increase understanding as to who responds favourably to these types of exercise.

An effective and economic method of performing this research could be to pool data from many exercise studies in MS. This would provide a large, multicentre sample, potentially drawing data from both European and American trials to improve the ecological validity of the findings. Limitations to this may be in the consistent use of outcome measures across trials for comparing results.

9.2 Minimising variability

The range of walking ability encompassed by EDSS 4.0-6.5 is great. In order to create a more homogenous sample future studies could stratify randomisation either by EDSS, baseline scores or type of MS. Potentially stratifying samples by symptomotology (for example a relevant primary MS symptom would be motor weakness) could potentially further assist predicting response to exercise.

9.3 Combined interventions

The SE intervention resulted in mean group changes in walking velocity that were clinically significant (21.7%) in this sample of ambulant people with MS. Larger studies have demonstrated that a course of Fampridine, a drug which aims to improve walking in ambulant people, results in similar improvements in walking velocity (25·2%) (Goodman *et al* 2009). Future research could investigate whether combining drugs (such as Fampridine) with exercise interventions would result in improvements greater than either Fampridine or exercise in isolation, thus maximising the effect of both interventions.

9.4 Exercise for people with MS who are not ambulant (EDSS > 7)

Most of the existing research, including this clinical trial, has been performed in people with EDSS <7.0. It may be more challenging to design exercise programmes for the more severely disabled (Asano *et al* 2009). Considering that people with MS have similar life expectancy to other people it is important not to overlook the effect of exercise interventions in people with EDSS >7.0.

In this clinical trial some participants used a wheelchair on occasions, however, the results do not inform us of the response to the exercise interventions in wheelchair dependent people with MS. Currently there has been one published feasibility study investigating the use of Pilates in this population (van der Linden *et al* 2013), but the response to exercise in more disabled individuals with MS is little known. Recently Skjerbæk et al (2014) demonstrated the feasibility of people with more severe disability (6.5≤ EDSS ≤ 8.0) exercising using predominately upper body endurance training. There is a need for future studies to evaluate the effectiveness of this and other differing types of exercise in wheelchair dependent people.

Studying exercise interventions early after diagnosis when disability is minimal is equally important, to determine its potential role in preventing progression. While there is some evidence to support the possibility of a disease-modifying potential of exercise (or physical activity) in MS patients, future studies using better methodologies are needed to confirm this (Dalgas and Stenager 2012).

The mean age of diagnosis of MS is approximately 35 years (Alcalde-Cabero *et al* 2013), and so long term adherence to exercise could be key to managing walking and balance impairments (Rietberg *et al* 2004). The challenge to clinicians and researchers lies in offering tailor designed exercise programmes that minimise the barriers to exercise (Garber *et al* 2011). In this clinical trial attendance at Pilates face to face sessions was lower than those attending the SE face to face sessions. Qualitative research could be performed to explore the reasons for this, either by contacting participants of this clinical trial or by

performing a new study. Identifying reasons which hinder people with MS from exercising is essential to providing appropriate and useful therapy services.

9.5 Methodology

The responsiveness of the secondary outcome measures; forward and lateral Functional Reach Tests, the Activities Balance Confidence scale and VAS have not been well established in the population used in this trial. Consequently, it was difficult to determine whether clinically significant changes had occurred in some of the measures. Studies to establish the SEM and MCID of the Functional Reach Tests in MS would be useful in interpreting the effects of interventions from existing data and would aid in calculations for determining sample sizes for future studies.

The data from this clinical trial could be used to further explore the validity of the Functional Reach Tests and Activities Balance Confidence Scale in MS. This would contribute to our understanding about objective outcome measures for use in MS research.

Section Two, Chapter One: The use of Ultrasound Imaging (USI) of the deep abdominal muscles.

Ultrasound imaging (USI) provides a method of visualising structures within the body in 'real time', which means that muscle activity can be imaged as it occurs (Hides et al 1998). This is useful for assessing current activity in muscles (Perkin et al 2003), and to analyse changes over time (Critchley et al 2011). USI is attractive as a method to measure change in muscles in both research and clinical practice due to its relative inexpense (in comparison with other imaging modalities) and ease of transportation (English et al 2012).

The deep abdominal muscles, namely TrAb and IO contribute to trunk musculature. Impairment in the trunk muscles is a common consequence of neurological pathology (Dickstein *et al* 2004) and can affect trunk stability (Lanzetta *et al* 2004) and balance (García-Vaquero *et al* 2012). Research demonstrates that delayed onset of activation occurs in TrAb of people with LBP (Hodges & Richardson 1999), however little is known about the behaviour of the deep abdominal muscles in people with MS. It is unclear whether MS impairs activation of these muscles. Additionally it is not known whether Pilates exercises, which aim to improve spinal stabilisation by voluntary activation of the deep abdominal muscles, result in changes in TrAb and IO in this population. To develop an understanding of this, USI was used as a method of measuring the deep abdominal muscles. It was considered that, in doing so, important information would be gained pertaining to the underlying mechanisms of change associated with the exercise interventions in this clinical trial.

This next section of the thesis will report and evaluate findings of the exploratory and experimental research performed to assess the deep abdominal muscles using USI. Firstly a reliability study was performed to assess the researcher (EF's) intra-rater reliability as an operator of USI and ascertain the stability of measurements. Additionally, it aimed to develop a protocol for use in the randomised clinical trial. Secondly a comparison of the USI measurements of people with MS with matched controls was performed. Thirdly the effect of exercises upon the deep abdominal muscles of MS was explored and finally post hoc correlations between the USI measurements and functional reach data from the clinical trial were evaluated. Alongside this a literature review was performed to report the psychometric properties of USI as a method for measuring the deep abdominal muscles.

It is worth noting at this point that multifidus is a paraspinal stabiliser which is often targeted in rehabilitation (Barr *et al* 2007) and can be measured using USI (Kiesel *et al* 2007; Koppenhaver *et al* 2009), however a pragmatic decision was made not to include measurements of multifidus in this study.

1.1 Psychometric properties of Ultrasound Imaging

Methods of imaging the deep abdominal muscles

In this section the psychometric properties of USI of the deep abdominal muscles will be discussed. USI allows a direct measurement muscle thickness changes and provides a convenient way of measuring muscle activity, atrophy and hypertrophy (Perkin et al 2003). However in order to be useful as a method of evaluating the effectiveness of interventions imaging is required to be

reliable, valid and responsive in detecting clinically significant changes (English et al, 2012).

Magnetic resonance imaging (MRI) or computerised tomography (CT) are considered to be the gold standard for measuring the size of skeletal muscles. Both are expensive modalities and CT scanning has the added complication of the risk incurred with ionising radiation (Pretorius and Keating 2008). Electromyography (EMG) can be performed to detect muscle activity using surface or invasive fine wire electrodes. Surface EMG is not useful for measuring activity in the deep abdominal muscles due to the depth of the muscles and the potential for 'cross talk' from adjacent muscles (Hides et al 1998; Hodges et al 2003). Fine wire EMG may be used to measure TrAb activity, however due to the invasive methodology, insertion of the wires may cause pain, bruising and fainting in participants (Hu *et al* 2011).

A disadvantage of USI as a measurement tool is that it is only able to image a 'slice' of the muscles directly beneath the transducer. Morphological differences have been identified between regions in the abdominal muscles that may reflect variations in functions (Urquhart et al 2005). Therefore muscle thickness changes seen on USI may not adequately represent the activity of all components of the imaged muscle (Hides *et al* 1998). However, this limitation would apply equally to EMG. Perkin et al (2003) proposed that the validity of USI as a measure of muscle activation magnitude is dependent on the muscle shape during contraction; for example the accuracy of measuring the external oblique (EO) muscle activity with USI has been questioned due to the change of shape (Brown and McGill 2010).

The reliability of USI in the measurement of TrAb and IO has been established (Koppenhaver *et al* 2009;Teyhen *et al* 2011) with intra-rater reliability considered to be more stable than inter-rater (Ferreira *et al* 2011; Teyhen *et al* 2011). Both the validity and reliability of USI will be explored in the literature review following. The results of the reliability study will be reported according to the guidelines proposed by Kottner *et al* (2011) .

1.2 Literature review

Search strategy

The search engines 'Embase' which includes Ovid Medline and PsycArticles, CINAHL and 'Google Scholar' were searched from 1974 to 17th November 2014. In order to focus specifically on the evidence pertaining to the validity and reliability of USI the following search terms were used as key words or words in the title/ abstract:- .

1. 'validity' and 'ultrasound' or 'sonography'
2. 'validity' and 'ultrasound' and ' lateral abdominal wall'
3. 'validity' and 'ultrasound' and ' deep abdominal muscles'
4. 'validity' and 'ultrasound' and ' Transversus Abdominis'
5. 'validity' and 'ultrasound' and 'Internal Oblique'
6. 'reliability and 'ultrasound' or 'sonography'
7. 'reliability' and 'ultrasound' and ' lateral abdominal wall'
8. 'reliability' and 'ultrasound' and ' deep abdominal muscles'
9. 'reliability' and 'ultrasound' and ' Transversus Abdominis'
10. 'reliability' and 'ultrasound' and 'Internal Oblique'

The results were sorted by relevance to include papers which measured reliability and / or validity of US measurements of TrAb and IO. EO was not included for two reasons; (1) EO was not included in the measurements taken of participants in this clinical trial, and (2) USI lacks accuracy in the reliability of measures due to the substantial changes of shape of the muscle during contraction. In addition a manual search was performed of the reference lists of the systematic reviews included.

1.3 Reliability

Valid inferences can only be made when instruments provide scientifically sound measurements. Ascertaining the reliability of USI measurements is important to ensure that any changes reported are not due to measurement error. An acceptable level of reliability depends upon the purpose of the test and should be predetermined prior to reporting. Throughout this document the ICC and confidence intervals (CI) will be reported and inferences may be drawn from table 25. The reliability of USI has been investigated by various authors and is dependent upon many factors. Multiple sources of error can affect the reliability of USI, in particular, if thickness changes are measured (see table 26).

A systematic review by Costa et al (2009) assessed the reliability of USI for the measurement of abdominal muscle activity. Twenty studies were included. Conclusions drawn suggest that the methodological design of the studies were suboptimal, making it difficult to establish the reliability of USI. Further research has been performed since then to establish criteria for the reliability of USI which are detailed below.

Classification	Intra-class coefficient (ICC)
Unreliable, inadequate for use, poor reliability	< 0.80
Adequate	0.80-0.90
Acceptable level for use, highly reliable	>0.90

Table 25: Classification of reliability

(adapted from Kottner et al 2011)

Potential source of error	Studies
Impact of visceral structures such as a full bladder which may compress upon TrAb	Teyhen et al (2007)
Contraction of an adjacent muscle such as EO compressing	Teyhen et al (2007)
Inaccurate identification of land marks	Ferreira et al (2011)
Position of subject and/or transducer	Ishida et al (2012)
Variation of performance of the activation task	Koppenhaver et al (2009)
Training of the operator	Teyhen et al (2011) Costa et al (2009) Ferreira et al (2011)
Food consumed	Kordi et al (2011)
Fatty infiltration to muscle from obesity and disuse atrophy	Thoirs & English (2009)

Table 26: Sources of measurement error in ultrasound imaging of Transversus Abdominis (TrAb) and Internal Oblique (IO).

Factors impacting on reliability:

Training requirements for operators

The ability and training of the operator can affect the reliability of USI. Teyhen et al (2007) and Ishida and Watanabe (2012) discussed the importance of diligent attention to steadying the position, orientation and inward pressure of the US transducer. Reducing medio-lateral transducer movement is important to ensure that any movement captured is due to activation of the muscles rather than movement of the transducer. Dupont et al (2001) reported that measurements of muscle thickness can be reduced by as much as 50% when strong contact pressure is applied, furthermore, angling the transducer away from a perpendicular approach reduced accuracy in distinguishing fascial planes.

Adequate training of the operator helps to improve technique and allow consistent inward pressure. Using a foam cube surrounding the transducer to help control movement is one method of assisting in the acquisition of reliable images (Ferreira *et al* 2011). Teyhen et al (2011) performed an inter-rater reliability study of USI of trunk musculature on asymptomatic soldiers (n=21). USI was conducted by novice operators who had undergone a 20 hour training programme. Automatic activation strategies were adopted using an ASLR . Inter-rater reliability is reported in table 27 and demonstrates that operators with 20 hours training are able to acquire reliable images of the deep abdominal muscles. These strategies for improving the quality of image acquisition are pertinent as the majority of error in measurement occurs whilst acquiring images and very little measurement error occurs when measuring the images on screen (Gnat *et al* 2012).

Muscle and position	Intra Class Co-efficient (95% Confidence Intervals)
Transversus Abdominis	
Rest	0.86 (0.65-0.94)
Active straight leg raise	0.87 (0.67-0.95)
Internal Oblique	
Rest	0.91 (0.77-0.96)
Active straight leg raise	0.93 (0.82-0.98)

Table 27: Inter-rater reliability of novice ultra-sound operators

(adapted from Teyhen et al 2011)

Koppenhaver *et al (* 2009) assessed the inter and intra-rater reliability of USI of TrAb and multifidus in subjects with LBP (n=30) using both an active straight leg raise (ASLR) and abdominal drawing in manoeuvre **(**ADIM) at end of expiration. Intra-rater reliability was reported ICC 0.93-0.98 (see table 28 page 240). Using the mean of two measures increased the reliability and precision of measurements. Conclusions drawn from this study suggest that intra-rater reliability is high for measuring the thickness of TrAb. It is noteworthy that the USI operators had received 70 hours of training in musculoskeletal USI, which is three times the amount received by those in Teyhen et al's (2011) study. Further to this recommendations proposed from a reliability study by Gnat et al (2012) are that the time between measurements did not significantly influence reliability (up to five days) and taking a mean of three thickness measurements

improves reliability. The between-days ICC is lower (less reliable) when measures are taken between-days rather than within-days, it is however still considered highly reliable.

Transversus Abdominis	**Intra-class Correlation coefficient (with 95% confidence intervals)**
Within-day	
Rest supine	0.98 (0.95-0.99)
Active Straight Leg Raise	0.96 (0.92-0.98)
Between-days	
Rest Supine	0.94 (0.87-0.97)
Active Straight Led Raise	0.93 (0.87-0.97)

Table 28: Reliability of same-day and between-day Transversus Abdominis measurements using ultrasound imaging.
(adapted from Koppenhaver et al 2009)

Ferreira et al (2011) assessed the level of operator training on reliability of US measurements. The trained operator received a three month training program in the US protocol and the non-trained received basic information on how to measure TrAb. Intra-rater reliability of thickness change was ICC 0.92, (CI 95% 0.81-0.87) for a trained operator and ICC 0.44 (CI 95% 0.41-0.78) for an untrained operator, using automatic activation strategies. This study further highlights that lack of operator training makes reliability of image acquisition inadequate. However the duration of training required for the operator to be

considered sufficiently trained is inconclusive, ranging between 20 hours and three months. Basic information on how to measure TrAb does not appear to be sufficient.

Additionally it is noteworthy that in taking between-day measurements food consumption prior to imaging may affect the reliability of results, as after a meal the thickness of TrAb has been found to decrease significantly (Kordi *et al* 2011).

Breathing mechanics

Breathing mechanics may affect reliability of TrAb and IO USI measurements. Kanaeko et al (2005) performed a study to assess reliability during both quiet breathing and forced expiration. Measures were taken at end inspiration and end expiration during quiet breathing, reliability was adequate (ICC 0.87-0.91). Significant increases in thickness of TrAb and IO were reported during forced expiration compared to ADIM ($p < 0.001$). Reliability of the measurements were varied (TrAb, ICC 0.66; IO, ICC 0.93, CI not reported) (note: this was an abstract, full paper in Japanese and not translated into English at time of writing). Recommendations by Kanaeko et al (2005) to maximise reliability indicate taking measurements at the same stage of the breathing cycle during quiet breathing. This was considered in the development of the protocol for our reliability study and our clinical trial.

Position of participant

Reeve and Dilley (2009) performed research into the effect of lumbar spine position on contraction of TrAb and found that erect sitting produced greater

contraction in TrAb than slumped sitting. A small change in the amount of lumbar flexion whilst sitting can effect the thickness of the deep abdominal muscles upon USI, this highlights the importance of clear, repeatable instructions when imaging participants in seated positions.

Norasteh et al (2007) assessed the effect of position on TrAb, IO, EO and RA with USI (n=27). Measurements were taken in three positions: standing, seated and supine. The exact positions of the participants were not reported. ICC for TrAb was 0.81 and IO 0.97 (CI not reported) for same day measures and 0.80 and 0.91 for TrAb and IO respectively one week apart. The authors did not report the reliability of the varying positions, the effect of expiration or any CI's which makes drawing conclusions from their work difficult, other than to say the reliability was adequate according to Kottner's guidelines (Kottner *et al* 2011).

1.3 Summary of reliability of ultrasound imaging for measuring Transversus Abdominis and Internal Oblique

To summarise, adequate reliability of USI to measure TrAb and IO has been confirmed by several studies, but may be affected by several factors. Operator training (and consequently error) has shown to play a significant role in the reliability. It is proposed that error is more likely to be made in the acquisition rather than the measurement of images; training therefore appears key to improving image acquisition. Factors which fall under this domain include: positioning of patient, pressure and position of transducer, taking images at the same point of the breathing cycle, method of acquiring the image (automatic or voluntary activation).

1.4 Validity: studies assessing ultrasound imaging in healthy people

Systematic reviews

The validity of a measure is the extent to which a test measures that which it is intended to measure (see, page 104) (Petrie and Sabin 2009). A systematic review by Perkin et al (2003) reviewed the validity and reliability of USI as an objective measurement of skeletal muscle activity for use within physiotherapy practice. The 11 studies reviewed correlated USI with either MRI, EMG or CT scans of a variety of skeletal muscles. All provided evidence to demonstrate that USI is a reliable and valid tool in healthy people and people with LBP, however not in an obese population. This review highlighted a number of factors which may affect validity and reliability which included: fat, fascial orientation, muscle shape (as described by Brown and McGill 2010, for EO) and pathology. They concluded there was a need for further research to evaluate USI as a tool for imaging deep and irregular muscles and pathology, since none of the studies investigated the validity of USI for measuring TrAb or IO.

A more recent systematic review by Pretorius & Keating (2008) assessed the validity of USI for measuring skeletal muscle size in comparison to a reference standard such as CT or MRI. To be included, all studies had to report a correlation co-efficient. All those correlating EMG with USI were excluded (the reference standard EMG did not measure muscle size). All seven studies demonstrated that USI is a valid measure of skeletal muscle size. However, this systematic review only included one study which explored the validity of USI for measuring the size of TrAb in young healthy sportsmen (n =13), of which the thickness of TrAb and IO measured by MRI scans was correlated with USI (ICC 0.78-0.95) (Hides et al 2006). Furthermore, none of the studies included

measurements of IO, or were undertaken in people with neurological impairments.

Mechanistic studies: healthy people

Initial studies were undertaken by Hodges et al (2003) to explore the validity of USI by correlating it with EMG. TrAb, IO and EO (n=3 healthy males) amongst other limb muscles such as biceps and tibilias anterior were imaged at contractions graded from 0-100% of maximal voluntary contraction (MVC) as determined by fine wire EMG recordings. Hodges et al concluded that USI can reliably detect 12% MVC for TrAb and 22% MVC for IO. The specific USI protocol was not detailed and the sample size was small, which makes drawing definite conclusions from this study difficult.

Hodges et al's findings were however supported by McMeeken et al (2004), who assessed the relationship between fine wire EMG activity and thickness change in TrAb (but not IO or EO) in healthy subjects, using 'abdominal drawing in'. TrAb thickness changes were very strongly correlated with EMG activity (R^2=0.87, p=0.001). Results demonstrated a linear relationship between EMG activity and TrAb thickness upon USI at all levels of contraction (McMeeken *et al* 2004).

The validity of using USI to measure abdominal oblique muscles is not as clear cut. Brown & McGill (2010) compared EMG and USI of EO and IO in healthy (n=5) males aged 25 (±3.8) years using both the ADIM and a full abdominal brace. Findings indicate that there was, at best, a weak relationship between USI thickness changes on EO and EMG activation levels on both abdominal

bracing and ADIM (EO, r= -0.22, 95% CI 0.42-0.01; IO r=0.14, 95% CI -0.09-0.35). Drawing upon these results, there are three points worth noting; firstly, in this clinical trial (my research), we did not measure EO with USI due to the uncertainties in relation to validity and reliability. Secondly, TrAb was not measured by Brown & McGill, hence no conclusions can be drawn from this study regarding the validity of USI for measuring TrAb. Thirdly, the ADIM is used to preferentially activate TrAb, not IO. The ADIM has been shown to activate TrAb 70% more than IO (Urquhart et al 2005). Hence it is unsurprising that the thickness changes demonstrated did not correspond to EMG activity. It may be that using automatic activation strategies for IO carry more validity.

In summary, the systematic reviews and mechanistic studies reported here support the use of USI as a valid tool for measuring TrAb activity in healthy people. However it cannot be disregarded that there has been more research performed in validity of USI to measure TrAb than IO, making it less certain as to whether USI should be used to measure IO.

Mechanistic studies of neurological populations

Perkin et al (2003) proposed that the validity of specific populations should be established before using USI as a research tool. Preliminary research to assess the validity of USI has been performed in people with acute stroke, using EMG as the gold standard measure (Hough et al 2009). EMG activity of TrAb and IO were simultaneously recorded with USI, using either hip flexion in supine or arm abduction in sitting (n=10) to automatically activate these deep abdominal muscles. The correlation for mean EMG recording and mean percentage thickness change for TrAb was r^2=0.62 and IO r^2=0.55 indicating that

percentage thickness change of TrAb and IO may be a valid measure of activity in people with stroke. To date USI imaging, however, has not been validated for use in people with MS. None of the other studies identified validated USI for use in neurological populations.

1.5 Summary of validity of ultrasound imaging for measuring Transversus Abdominis and Internal Oblique

In summary, there is a reasonable body of evidence to support the use of USI as a valid tool for measuring the cross sectional thickness and activity in TrAb in both healthy populations, and some evidence to support its use in neurological conditions such as stroke. There is less evidence to support the use of USI for measuring IO. The discrepancies in correlating USI with EMG for IO may lie in the method of activation. Automatic activation of this muscle appears to correlate better with EMG when using ADIM. The presence of fatty infiltration to the muscle or deep adipose tissue can impact negatively on the validity of USI as a measure of the deep abdominal muscles. Furthermore USI has not proven valid for measuring EO.

1.6 Responsiveness of ultrasound imaging measurements of Transversus Abdominis and Internal Oblique

Responsiveness has been defined as the power of a measure to detect a clinically significant change, that is a change beyond measurement error (see chapter four, page 105) (Guyatt *et al* 2002). Whilst there has been sufficient evidence published to determine that USI is a reliable and valid tool for the measurement of deep abdominal muscles (Koppenhaver *et al* 2009; Ferreira *et al* 2011; Teyhen *et al* 2011; English *et al* 2012), few studies have focused upon

establishing responsiveness and furthermore MCID. However, those which have measured the thickness of TrAb and IO have included the SEM and MDC (Gnat *et al* 2012), which can be used as reference points. From this, one can gauge whether changes are beyond measurement error. Currently, there has been no conclusive evidence to define a clinically significant improvement in any population.

Table 29 details the SEM and MDC drawn from reliability studies. The SEM for TrAb USI measurements taken at rest in a supine position lie between 0.1mm - 0.48 with the MDC between 0.4mm -1.34mm. Only one study calculated these for IO (table 29). The MCID has been reported as 1.77mm for TrAb and 2.15mm for IO (note this was abstract only, full text was in Japanese so unable to determine how this was calculated). The clinical relevance of these figures has yet to be established, however Koppenhaver et al (2009) suggest that the percentage increase of TrAb would need to be 133% for the clinician to be 95% confident that change was beyond measurement error, with slightly lower changes reported by Teyhen et al (2011) and Gnat et al (2012). This may be attributed to differing methodologies; Teyhen et al used an ASLR and Gnat et al used ADIM. Using an ADIM may result in a higher MDC due to the voluntary control required to modulate abdominal muscle contraction.

Taking measurements one week apart increased the MDC by 8.7% as reported by Koppenhaver et al(2009). This could give rise to questions regarding the stability of measures taken before and after intervention periods, such as the 12 week intervention period for this study. This is relevant as intervention periods

of greater than eight weeks are required produce hypertrophic changes in muscle tissue (Danneels *et al* 2001; Dorado *et al* 2012).

In summary, there is a small body of research evidence which details the SEM and MDC of US measurement of TrAb, however the clinical significance of changes occurring in the deep abdominal muscles has not been established. In order to determine the clinical relevance of thickness changes, other factors such as the onset of muscle activation need to be considered (Vasseljen *et al* 2009). Section 2, Chapter 4 page 286 explores the relevance of thickness changes in the deep abdominal muscles in more detail.

Author	Muscle and position	SEM	MDC	Comments
Koppenhaver et al (2009)	TrAb supine rest	0.1mm (calculated from with-in day reliability measurements) 0.2mm (between- days)	0.4 mm within- day 0.6 between-days	A percentage change of 133% at ADIM would have to occur for the clinician to be 95% confident that change was beyond measurement error.
	TrAb ASLR	0.3 with-in day 0.4 between- day	0.8 within-day 1.1 between-days	
	Percentage change	9.2% within day 12.3% between day	25.4% within-day 34.1% between day	
Teyhen et al (2011)	Supine TrAb at rest	0.4mm	1.0mm	SEM is approximately 7-10% of resting thickness and MDC is 20-25%
	Supine IO at rest	0.7mm	1.9mm	
	TrAb ASLR	0.4mm	1.1mm	
	IO ASLR	0.7mm	1.9mm	
Gnat et al (2012)	TrAb at rest supine	0.18-0.48mm	0.60-1.34mm	Higher reliability with inter-rater and between day
	TrAb supine ADIM	7.28%-18.91%	20.18-53.43%	
Arab et al (2013)	TrAb at rest supine	0.19mm with-in day 0.21 between-days	0.52mm within- day 0.58mm between-days	SEM and MDC are higher for all abdominal muscles in a clinical population (not reported here) and when measured in unstable postures i.e. sitting on gym ball.
	IO at rest supine	0.20mm both with-in and between-days	0.55mm within- and between-days	
Yang & Park (2014)	TrAb rest	0.13mm	Not reported	
	IO rest	0.16mm	Not reported	

Table 29: Ultrasound measurements: Standard error of measurement and minimal detectable changes of Transversus and Internal Oblique in healthy populations.

1.8 Methods of activating the deep abdominal muscles

There are various methods to activate the deep abdominal muscles. Voluntary activation has been described, which can be achieved with the use of an ADIM (Urquhart et al 2005) or a full abdominal brace, in which there is forced contraction of the entire abdominal wall (Brown and McGill 2010). Automatic activation of the deep abdominal muscles can be achieved using the ASLR (Hu et al 2002) or by limb flexion (such as contralateral shoulder movement) (Hodges *et al* 2003) or supported hip flexion (Ferreira *et al* 2011).

Pilot work

Pilot work was performed on staff volunteers to practise technique and establish a protocol. EF performed approximately of 20 hours training (partly supervised and partly practising unsupervised) from an experienced researcher in the field of USI (Dr Alan Hough) in line with recommendations (Teyhen *et al* 2011). Pilot work was performed to explore the intra-rater reliability of measuring automatic activation using contralateral arm lift activation positions which had shown promising validity compared to needle EMG (Hough et al 2009). Difficulty was experienced in reproducing and maintaining the sitting postures during this procedure which made reliable imaging more problematic; in particular participants tended to slump during the imaging procedure. It was also observed that in sitting participants tended to demonstrate some intermittent voluntarily activity even when not requested to. The researcher (EF) proposed that this may have been in response to the abdominal flesh being exposed, which resulted in them drawing in the abdominals. Upon examining the literature it was noted that the protocols of other studies did not include seated

measurements in reliability studies (Koppenhaver *et al* 2009; Ferreira *et al* 2011 ;Teyhen *et al* 2011) hence there was a paucity of literature to compare these findings to. Adaptations to the protocol were therefore made, to include a supine rest measure and an ASLR as an activation measure (as Teyhen et al 2011), as during the practice and training sessions the operator (EF) found that acquiring images of TrAb and IO with the subjects in a supine position resulted in clearer images of the fascial borders.

This next section will evaluate methods of activating the deep abdominal muscles and the biomechanics of ASLR as the chosen method of activating TrAb and IO in our reliability study and clinical trial.

Abdominal muscle activation increases lumbar spinal stability (Stokes et al 2011)**.** Urquhart et al (2005a) reported that a drawing in manoeuvre activated TrAb 70% more than IO, and 100% more than EO (p=0.001, n=7, asymptomatic people) measured with fine wire EMG. The relative effectiveness of this manoeuvre to increase the stability of the spine, in comparison to a full abdominal brace, in which all the abdominal muscles contract, has been questioned (Grenier and McGill 2007). Functionally, the spine requires flexibility in addition to stability and consequently a full abdominal brace increases spinal stiffness (Grenier and McGill 2007) due to the co-activation of all the abdominal muscles, whereas the ADIM preferentially activates TrAb.

The magnitude of activity (measured by fine wire EMG) in the deep abdominal muscles produced by an ADIM is subject to variation (Bjerkefors *et al* 2010) and may be dependent on an individual's effort and level of body awareness

(Urquhart et al 2005a). Reliability of USI measurements when performing ADIM were considerably lower than using an ASLR (ADIM ICC 0.56, 95% CI 0.30-0.74; ASLR ICC 0.98, 95% CI 0.97-0.99) (Gnat *et al* 2012), which could be attributed to difficulty in ascertaining the MVC during the ADIM. This research was performed in healthy people and the effect of neurological pathology on the ability to modulate deep abdominal muscle activity is not known. Limb movements destabilise the spine and consequently the deep abdominal muscles, namely TrAb and IO are recruited (Hu et al 2012). The ASLR is therefore not dependent on the participant having the ability and body awareness to voluntarily activate the abdominal muscles.

The ASLR can also be used a biomechanical test to assess load transfer between the legs and spine via the pelvis (Mens et al 2001). During the ASLR a subject lies supine with knees in extension and lifts one leg to 5- 20 cm (depending on author varying heights reported, hence detailing the height is crucial for repeatability) (Liebenson *et al* 2009). The hip flexor muscles (rectus femoris and illio-psoas) exert an anterior force in which the ipsilateral ilium is pulled into anterior rotation (Hu et al 2012). In response, TrAb, IO and EO activate to stabilise the pelvis in a form of forced closure of the sacroiliac joint, which further stabilises the lumbar-pelvic girdle (O'Sullivan et al 2002; Beales et al 2009). Whether these muscles act ipsilaterally (Hu et al 2012), contralaterally (Teyhen *et al* 2009) or bilaterally (Teyhen *et al* 2009) remains unclear. Research in this area is contradictory as these muscles have been shown to act both symmetrically and asymmetrically to assist in the stabilisation process (Tsao et al 2008).

Furthermore psoas has been advocated to act as a stabiliser of the lumbar spine during the ASLR (Hu *et al* 2011). Using fine wire EMG in healthy subjects (n=17) Hu et al showed that during an ASLR, ipsilaterally, psoas acted in conjunction with iliacus and rectus femoris as a hip flexor. Contralateral to the ASLR psoas was active indicating that it plays a role in supporting the anterior lumbar spine in response to destabilisation. As an explanation Hu et al suggested that psoas compresses the facet joints in the lumbar spine to aid stabilisation. Due to the invasive nature of measuring psoas activity there are few studies in which the role of this muscle is analysed except those undertaken by Hu et al (2011, 2012). The relevance of this lies in understanding that individuals employ differing neuromuscular strategies to stabilise the spine (Morris *et al* 2012) in which psoas may play a role. Although psoas is not visualised during USI of a cross section of the lateral abdominal wall, an awareness that it may contribute to stability assists in understanding activation patterns captured by USI.

With the objective of further understanding the mechanisms of muscle activity which occur during an ASLR, Hu et al (2012) undertook a study whereby surface EMG was applied to IO, EO, RA, rectus femoris (RF) and biceps femoris (BF). Fine wire EMG was inserted into TrAb and IO (n=16). Results suggest that muscle activity in TrAb, IO, and RF was significantly greater when the leg on the ipsilateral side was raised ($p<0.03$) and BF greater with the contralateral leg ($p<0.05$). With regard to symmetry, most participants demonstrated ipsilateral activity in TrAb and IO, (numbers not disclosed), although some adopted a different strategy. This further supports the notion that even in healthy people the neuromuscular strategies utilised are individual.

Additionally EO may contribute to spinal stability during the ASLR (Hubley-Kozey *et al* 2009).

Furthermore when participants performed the task repeatedly there were large variations in the magnitude of EMG activity, which may be accounted for by rehearsal or fatigue. Extrapolating from this research assisted in the development of a protocol for our reliability study and clinical trial. Also taken into account was the potential effect of MS upon the ability to ASLR and activate the deep abdominal muscles. Weakness arising as a consequence of MS often results in a participant having a stronger and weaker side. Participants were asked to perform the ASLR with their stronger leg. In line with recommendations by Gnat et al (2012) taking more than one image would also allow for rehearsal and fatigue. The mean was calculated from these and used for data analysis.

1.9 Biomechanics of the active straight leg raise in people with neurological conditions

There is a paucity of published research to determine the biomechanics of the ASLR in people with neurological conditions, with just one study performed. Gatti et al (2008) compared surface EMG of the quadriceps and RA of people with MS and healthy people (n=14 + 14 matched controls), stating that the aim was to assess the effect of MS upon muscles which stabilise the lower limb during an ASLR. Significant differences between groups were reported ($p=0.006$) for activity of the RF muscle, however they do not state which group had the greater degree of activity. Neither 95% CI or data for RA were reported for between group differences, which makes it difficult for the reader to draw

clear conclusions regarding potential differences in EMG activity between people with MS and healthy controls.

In summary very little is known about the behaviour of the abdominal muscles in people with MS. The effect of other neurological pathology on the deep abdominal muscles is further discussed on page 268.

1.10 Summary of literature surrounding activation of abdominal muscles during the active straight leg raise

The ASLR can be used as a method of destabilising the spine in order to activate and evaluate abdominal muscle activity. This has been studied in healthy people using EMG (Hu et al 2011 & 2012) and in people with LBP using USI (Teyhen et al 2009; Ferreira et al 2004). The research demonstrates that TrAb and IO muscles act to stabilise the lumbar spine in response to movement and that EO may also contribute to stability. The role of psoas as a stabiliser has been proposed by Hu et al (2012). There is currently little evidence about the behaviour of the stabilising muscles in people with neurological impairments, with only one study having been performed in people with MS (Gatti et al 2008).

The ASLR does appear to be an appropriate method of activating the deep abdominal muscles in order to measure them with USI. Despite the fact that individuals employ differing strategies for activation, unlike the ADIM, using an ASLR does not require the same levels of bodily awareness and abdominal muscle control which may make it a more appropriate technique to be used

where sensory and motor deficit may be present, as can be the case in people with MS.

Section Two, Chapter Two: Reliability study

2.1 Introduction

The aim of the reliability study was firstly to develop a protocol for USI for use in the clinical trial and secondly to ascertain the reliability of measurement of TrAb and IO by USI by the researcher (EF), in a healthy population prior to performing USI in the clinical trial.

2.2 Methods

Participants

A convenience sample of 10 healthy participants was recruited, via poster advertisements from Plymouth University and a local outdoor activities centre, for this reliability study (Ethical approval: trial registration: NCT01414725, IRAS 10/H0106/88). Participants were excluded if they had any history of LBP that restricted function or resulted in time off work, debilitating illness or pregnancy within the last two years in line with Ferreira et al (2011) and Teyhen et al (2011).

The sample size calculation was determined according to the minimum acceptable lower band 95% confidence interval of 0.8 as suggested by Donner and Eliasziw (1987) and further supported by Hobart *et al* (2012). Ten subjects were imaged on two occasions.

Ultrasonography

Real-time B-mode US images were acquired with a MySono U5' ultrasound system (Medison Ltd, Korea) and a wide-band linear array transducer with centre frequency 7.5Mhz. All images were taken by the same operator (EF).

Standardisation of imaging procedure

Participants were positioned supine on a plinth. The transducer head was placed transversely across the abdominal wall, midway between the inferior angle of the rib cage and the iliac crest, approximately 10 cm from mid line, as described by Norasteh et al (2007), Costa et al (2009) and Teyhen et al (2011). Any anatomical anomalies such as moles, tattoos etc. were documented to aid replacement of the transducer onto the same position on the second occasion. Three US cine-loop image clips were taken at rest during quiet respiration to capture abdominal muscle activity during at least one inspiratory and expiratory phase. A contra-lateral ASLR to 5cm off the plinth was then demonstrated to the participant (replicated as Teyhen et al 2011) by manually positioning the leg. The participant was then asked to repeat the ASLR three times whilst cine-loop image clips were taken. The participant was then positioned in unsupported sitting on a plinth with both hips and knees flexed to 90 degrees, feet level and supported. Participants were then instructed to 'sit up as tall as possible and then relax 10%'. Cine-loop image clips were taken first at rest during quiet respiration in this seated position and again during automatic activation while the contra -lateral arm lifted a weight chosen by the participant to be 'moderately difficult'. Weights offered ranged between 0.5-5 kg. Each of the measures was repeated three times as Hough et al (2009). This procedure was repeated on a second occasion two hours later.

Analyses of data

Captured cine-loop clips (avi format) were transferred to a PC for analysis. The cine-loop clips were converted to a sequence of bitmap images using Virtual Dub (Version 1.9.11, Avery Lee, 1998-2010) and then imported into Image J (Version 1.46r) for measurement of muscle thickness. Images were not blinded prior to measurement. For the resting sequences the images were viewed in Image J and the maximum and minimum cross sectional thickness, which was identified visually by watching the image clips to account for inspiration and expiration. The superficial and deep borders were represented by hypoechoic fascial lines (Teyhen et al 2011).

Cross sectional measures were taken of TrAb and IO (see figure 10). Three measurements were taken, one measurement was taken midline of the image and one cm either side, then the mean was calculated to increase the reliability (as Koppenhaver et al 2009 and Teyhen et al 2011). Measurements were performed offline using 'Image J 'image measurement processing software. The distance between the inner fascial borders of TrAb and IO were measured in pixels and converted to millimetres.

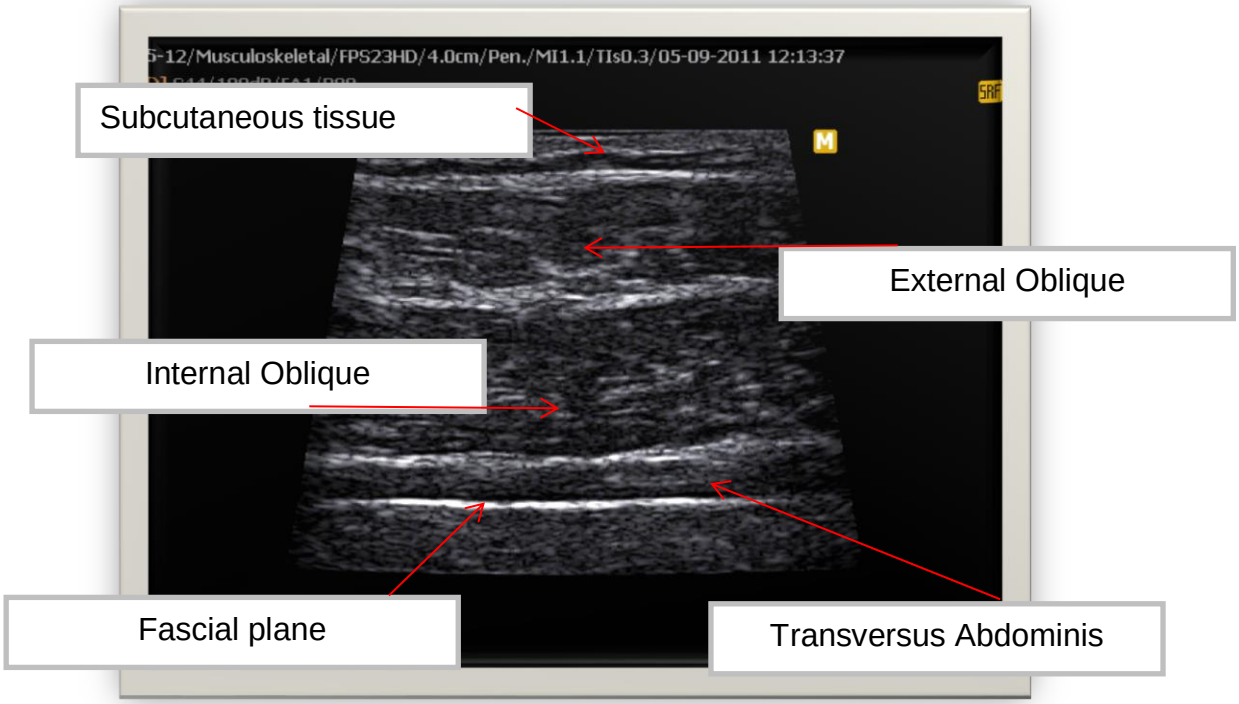


Figure 10: A resting US image of the deep abdominal muscles, taken by the researcher (EF)

2.3 Results

All 10 of the participants recruited were included in the data analysis.

Participants had a mean age of 35.8 years (range 26-64 years) with a ratio of 6:4 males to females. The thickness of TrAb and IO during rest and activation are reported in table 30 and the reliability measures (ICC, CI and SEM) are reported in table 31. The supine measurements for TrAb and IO were considered highly reliable (ICC> 0.90). The seated measurements taken for seated activation were not adequately reliable (ICC<0.80).

Test position	TrAb (Mean ± SD mm)		IO (Mean ± SD mm)	
	1st occasion	2nd occasion	1st occasion	2nd occasion
Supine rest	3.5±1.1	3.4±0.9	8.1±2.8	8.3± 2.4
Supine ASLR	4.3±1.2	4.2±1.2	8.3±3.3	9.1±3.5
Sitting rest	5.5 ±2.3	4.8± 1.2	8.3 ±3.3	9.1± 3.5
Sitting arm lift	6.3± 3.0	5.3± 1.4	12.0 ±3.4	12.3± 3.6

Table 30: Summary of measurements for Transversus Abdominus (TrAb) and Internal Oblique (IO) thickness on two repeat occasions (n=10)

Test position	TrAb (Mean ± SD mm)		IO (Mean ± SD mm)	
	ICC (95% CI)	SEM (mm)	ICC (95% CI)	SEM (mm)
Supine rest	0.98 (0.92-0.99)	0.21	0.98 (0.89-1.00)	0.54
Supine ASLR	0.99 (0.94-1.00)	0.23	0.98 (0.93-1.00)	0.53
Sitting rest	0.96 (0.86-0.91)	1.63	0.96(0.86-0.99)	0.91
Sitting arm Lift	0.36 (-0.31- 0.79	2.09	0.71 (0.19-0.92)	2.67

ICC = Intra class correlation coefficient (3,k); SEM =Standard error of measurement

Table 31: Reliability results for ultrasound measures of Transversus Abdominus (TrAb) and Internal Oblique (IO) thickness

2.4 Discussion

Summary of results

In this study intra-rater reliability measurements of supine resting and activation of TrAb and IO and seated resting of TrAb and IO were determined as 'adequate' for research purposes as defined by Kottner *et al* (2011) see table 25 page 237. However intra-rater reliability measurements of seated activation of TrAb were poor. While the ICC of 0.71 for IO might at first appear to be

adequate, the wide confidence intervals mean that this measurement approach, in its current format, is not sufficiently reliable for use in research.

Reliability levels of the researcher (EF) compare favourably with published work in the field (see table 32). It is noteworthy that studies may overestimate reliability if the raters are not blinded to previously obtained values. Further to this reliability is more trustworthy if 95% confidence intervals are reported (Kottner *et al* 2011; English et al 2012).

	Reliability Intra-rater ICC	**Koppenhaver et al 2009 Intra-rater ICC**	**Teyhen et al 2011 Inter-rater ICC**	**Ferreira et al 2011 Intra-rater ICC**
TrAb rest (supine)	0.98	0.98	0.86	Not reported
TrAb ASLR	0.99	0.96	0.87	0.92
IO rest (supine)	0.99	Not reported	0.91	Not reported
IO ASLR	0.98	Not reported	0.93	Not reported
ASLR= Active Straight Leg Raise, ICC= Intraclass Correlation Coefficient , IO= Internal Oblique, TrAb= Transversus Abdominus				

Table 32: comparison of intra-rater ICC measures with published data

Potential sources of measurement error

Reasons for the variation in reliability vary. Operator skill and training has demonstrated to be of great importance for acquiring reliable images (Ferreira et al, 2011; Teyhen et al, 2011; Koppenhaver et al, 2009; Gnat et al 2012). The

researcher (EF) performed in the region of 20 hours training, considerably less than the three months training that Ferreira et al received, which may have given rise to errors in image acquisition. Error may have occurred due to slipping of the transducer against the skin as the abdominals contract, resulting in the central part of the muscle not being imaged. Teyhen et al (2011) recommended using two hands on the transducer in addition to a high density foam cube. Keeping two hands on the transducer was not possible due to the lack of a foot control switch or assistance of another person in this study. Teyhen et al (2007) also suggested that the use of a foam cube may be inhibitory to the participant performing certain tasks, however, this may not be the case in performing a straight leg raise, as was used in the final protocol. Attempts at reducing aspects of operator error were made by accurate documentation of the transducer position in the participant's records; with anatomical mapping of the bony land marks, skin anomalies, moles, tattoos etc.

Variability in the pressure of probe placement may have also impacted on reliability. Dupont et al (2001) reported that downwards pressure on the transducer may reduce the cross sectional measurement of a muscle by 50%. Prior to undertaking the reliability study investigations were made into placing a pressure gauge on the US transducer probe to standardise pressure but this was rejected as it was decided that this would obscure the images taken. Using a foam cube to stabilise pressure as Ferreira et al (2011) may have prevented human error with regard to this.

In measuring the images error may have occurred by the blurring of the fascial borders of IO and TrAb. As suggested by Teyhen et al (2011) the medial

borders were taken as a point of measurement. Good image acquisition results in images with sharp fascial planes with high echogenicity. Contraction of TrAb and IO resulted in the fascial borders blurring on some images and making it difficult to accurately define a specific measuring point. In the case of blurred borders an estimate was taken at a mid-point. This was a pragmatic decision which was discussed in a supervision session with one of the supervisors (AH).

The morphology of IO may have contributed to reduced reliability of seated measurement. The shape of IO is similar to a crescent when contracted (Urquhart et al, 2005, Teyhen et al, 2007) and there was sometimes difficulty in ascertaining the mid-point. In contrast TrAb remained quite uniform in shape upon contraction. Error was reduced by taking the mean of three measurements as recommended by Koppenhaver et al (2009) and Gnat et al (2012).

Activation patterns

A qualitative observation whilst watching the cine loop clips (a four second period captured on film) was that upon automatic activation of both TrAb and IO, the muscles immediately increase in size, then fluctuate before stabilising back to the same thickness observed immediately after activation. Measurements were taken at the thinnest and thickest point of the muscle to account for fluctuations and respiration. The thinnest/ thickest point was determined by a brief visual analysis. In our study this phenomenon was observed in both very fit, Pilates trained individuals and less active individuals. To date this exact phenomenon has not been reported in the literature. One explanation for this may be an anticipatory feed forward reaction occurs in order to stabilise the lumbar spine in preparation for the straight leg raise. The

anticipatory feed forward reactions of TrAb have been reported in response to arm movements (Hodges & Richardson 1999). Individual patterns of onset within the abdominal muscles vary (Allison and Morris 2008). Further investigation, using cine-loop video to observe and quantify this observation would be useful to determine whether measurement error could occur as a consequence of this pattern of muscle activation.

Posture has great influence on variations of TrAb thickness (Reeve and Dilley, 2009). Research demonstrates a linear relationship between erect sitting (lumbar spine neutral), slumped sitting (lumbar flexion) and supine lying. TrAb are thickest during erect sitting (Reeve and Dilley 2009 ; Rasouli et al 2011). The intra-rater reliability of seated measures of both IO and TrAb thickness were poor in this reliability study. Difficulty in reproducing exact sitting postures may have resulted in the varying degrees of activation and therefore thickness.

Setts et al (2009) demonstrated that voluntary drawing in manoeuvres produces the greatest percentage change in thickness of TrAb. Despite instructing participants to relax during sitting, participants with adipose tissue over the abdominal muscles may have unconsciously drawn in the abdominal muscles as they felt self-conscious about the abdominal flesh being exposed. Although not supported by research, participants reported this during USI in this study.

Visual examination of the US images highlighted between-participant differences in deep abdominal muscle recruitment strategies during activities. Visual qualitative analysis suggested that IO was the more dominant stabilising muscle in some people. This observation is congruent with research suggesting

that individuals differ in strategies of activation (Allison & Morris 2008; Westad et al 2010; Hu et al 2012)

The results of this study indicate that intra-rater reliability of a trained operator was adequate for research for supine resting and ASLR measures for TrAb and IO, but inadequate for seated resting and activation measures. In light of this the protocol for the image acquisition in participants with MS was refined to include only supine resting and automatic activation measures. A limitation of this reliability study was that healthy, and mainly young subjects were used, not people with MS. The deep abdominal muscles of people with neurological pathology may behave differently and US images may differ due to disuse atrophy or neurophysiological changes (Perkin *et al* 2003).

2.5 Conclusions of the reliability study

USI has been reported by several authors as a reliable and valid measure of thickness of TrAb and IO. Intra-rater tends to be greater than inter-rater reliability and is highly dependent operator training. Image acquisition is influenced by a number of factors including subject and transducer positioning and activation strategy. The results of this reliability study demonstrate that the researcher (EF) is highly reliable in acquiring and measuring US images taken when the participant undertakes an ASLR in supine, but is unreliable for images acquired during seated automatic activation.

Section Two, Chapter Three: Ultrasound Imaging of the deep abdominal muscles of people with MS during automatic activation: a comparison with matched controls

3.1 Introduction

USI has been used to assess changes in the deep abdominal muscles resulting from chronic LBP. Reduced thickness increases upon activation have been reported for TrAb in comparison to healthy controls with back pain (Critchley & Coutts 2002; Ferreira et al 2011;Teyhen et al 2009) and in amputees (Springer & Gill 2007), suggesting that USI can be used to detect changes in response to pathology or injury. The published research has mainly focused upon USI of the deep abdominal muscles in the presence of LBP. Due to the lack of published research in MS, and in order to provide a detailed understanding as to the role of the deep abdominal muscles and the effects of pathology, this chapter will also consider populations other than those with MS and methods of measurement such as EMG in the literature review. The aim of this study was to use USI to measure automatic activation of the deep abdominal muscles, namely TrAb and IO, during an ASLR in people with MS and to compare this with matched controls.

3.2 Literature Review

Search strategy

This literature review examines published evidence which investigates the relationship between MS and USI of the abdominal muscles. In order to focus specifically on the evidence pertaining to the evaluation of the deep abdominal muscles in people with MS the following search strategy was used. The search

engines ‘PubMed and ‘Embase’, which included OvidMedline and PsycArticles were searched from 1974 to 17th November 2014. The inclusion criteria were: written in English, published in peer reviewed journals, and using EMG or USI to evaluate the role of the abdominal muscles. In addition, a manual search was performed by screening the reference list of key papers. To assess methodological quality guidance was sought from Greenhalgh (2008). As studies performed in this area tend to be small mechanistic studies CONSORT, PEDRO and TiDIER guidelines were not used. The search generated numerous articles which focussed on reliability of USI; these have been discussed in on page 236-242, and will not be critiqued again in this chapter.

Search terms

1. ‘Multiple sclerosis’ or ‘stroke’ or ‘neurological’ and ‘ultrasound’ and ‘lateral abdominal wall’, n = 7
2. ‘Multiple sclerosis’ or ‘stroke’ or ‘neurological’ and ‘ultrasound’ and ‘abdominal muscle’, n = 1
3. ‘Multiple sclerosis’ and ‘internal oblique’, n=0
4. ‘Multiple sclerosis’ and ‘transversus abdominis’, n =0
5. ‘Multiple sclerosis’ or ‘neurological’ or ‘stroke’ and ‘electromyography’ and ‘abdominal muscle’, n= 23

Activation of the abdominal muscles: Comparing neurological conditions with healthy controls.

Two relevant studies investigating deep abdominal activity in neurological populations were identified. Gatti et al (2008) (study described on page 254)

proposed a difference in the quality of activation in people with MS in comparison to healthy controls, although how they concluded this is unclear.

Postural muscle has been considered to require low levels of activity (one to three per cent of maximum voluntary contraction) in order to stabilise the spine during unsupported postures (Cholewicki 1997). Inactivity such as prolonged bed rest requires little activity the TrAb muscle, and has been found to result in decreased thickness upon USI (Ikezoe et al 2012). Ikezoe et al used USI of the TrAb and IO to assess potential thickness changes of these muscles between three groups; young active (n=11, 20 years± 0.8), elderly active (n= 28, 85.7 years± 5.5) and elderly prolonged bed rest (n= 13, 87.9 years ± 6.3). Significant differences were reported for the resting thickness of TrAb of the elderly bedridden group in comparison with the active groups. This indicates that when postural muscles are not activated for prolonged periods, atrophic changes may occur. However the study sample was free of neurological pathology hence assumptions cannot be drawn regarding the effect of MS on the deep abdominal muscles. Furthermore the sample used for our study was ambulant (EDSS 4-6.5) so not comparable to those on prolonged bed rest.

Unilateral stroke can result in reduced trunk muscle strength (Karatas *et al* 2004). EMG activity of superficial trunk muscles (RA and ES) is reduced post stroke (Dickstein *et al* 1999). More recent research has investigated the effects of movements which destabilise the spine (head lift and hip flexion)upon activity in the deep abdominal muscles post stroke (n=11+11 matched controls) (Marsden *et al* 2013). The study used fine wire US guided EMG of TrAb and IO to detect activity. Interestingly the findings reported no significant difference in

the magnitude of activation in either ipsilateral or contralateral sides (to the stroke) in either groups (stroke or control) during a head lift. Explanations provided for this pertain to the symmetrical activation of the deep abdominal muscles resulting from bilateral projections from the motor cortex and brain stem to the trunk muscles (Kuypers & Brinkman 1970; Murayama et al 2001).

The research performed by Marsden et al (2013) raises questions regarding how much of an effect MS might have on the deep abdominal muscles. A delay in anticipatory muscle onset in response to lifting a hand weight in RA and ES muscles has been demonstrated in people with MS in comparison to matched controls, even in mildly affected people (Krishnan *et al* 2012). Krishnan et al, measured paraspinal but not deep abdominal muscles, Murayama et al (2001) reported that paraspinal muscles receive contralateral cortical innervation suggesting that cortical lesions may have a greater effect upon the paraspinal muscle activity than on deep abdominal muscle activity, which would explain why delayed muscle onset might occur in those people with MS with cortical lesions. Additionally Krishnan et al (2012) measured *onset* of activation whereas Marsden et al (2013) measured magnitude of activation.

Furthermore, Marsden et al (2013) found that in response to hip flexion, bilateral activation of the deep abdominal muscles was higher when moving the paretic leg (in the people with stroke). In contrast, matched controls demonstrated a greater level of activity during ipsilateral hip flexion, suggesting that deep abdominal muscle activity in stroke may be a compensatory activity and/ or 'overflow' which is used to stabilise the spine in the presence of neurological weakness.

The dearth of published scientific literature relating to the contribution of the deep abdominal muscles to spinal and consequent trunk stability in people with MS makes this area poorly understood. Historically, exercises to stabilise the trunk have been employed by therapists working in this area (Smedal *et al* 2006) and furthermore Pilates exercises are used in clinical practice with the aim of activating the deep abdominal muscles, which has largely been based on a theoretic rationale (Freeman *et al* 2012). The purpose of this study was to aid the understanding of any existing differences between TrAb and IO in people with MS and matched controls by comparing US images of activation of these muscles during an ASLR.

3.2 Methods

Recruitment and eligibility criteria

Twenty people with MS were recruited via the SWIMS database newsletter. All methodology regarding recruitment, inclusion and exclusion criteria is described in detail in the chapter detailing the methods of clinical trial, page 141.

Of the US images taken for the clinical trial, 17 were of sufficient quality to use therefore 17 matched control participants were recruited via poster advertisement at Plymouth University and the University of the Third Age. A sample size calculation was not performed as the research was exploratory in nature. Control participants were matched to people with MS by gender, age (+/- five years), and by visual assessment of body frame. All control participants were free from neurological, cardio-respiratory or musculoskeletal pathology which could affect the trunk muscles, LBP (within the last three months) and

were not pregnant. Height and weight was recorded for both groups. Ethical approval was gained, as part of the main clinical trial, from the National Research Ethics Service, South West 3 Regional Ethics Committee (REC Reference Number: 10/H0106/88), and from the Faculty of Health and Human Sciences Ethics Committee at Plymouth University (REC Reference Code: MS/ab).

Procedure

USI was performed after written consent was taken, in line with the procedure described on page 258.

Statistical analysis

Demographic data was summarised using descriptive statistics. The average (mean) thickness and standard deviations were calculated using Microsoft Excel 2012. Thickness changes were expressed as a percentage, which were calculated from the equation (Teyhen *et al* 2009):

$$\text{Percentage change} = \left(\frac{\text{activation thickness- resting thickness}}{\text{resting thickness}}\right) \times 100$$

Significance level was set at p=0.05. Data was tested for normality using a Kolmogorov-Simonov test in SPSS. A sample of raw data was quality assured by the director of studies (JF) and academic supervisor (AH). Any outliers in measurements of raw data (US scans) were re-measured. In the case of unclear images advice was sought from the academic supervisor (AH) Outliers were identified as values greater or less than the mean ± two standard deviations and were replaced by the mean ± two standard deviations as

directed by Field (2009 page 153). Differences in TrAb and IO thickness at rest and during activation were analysed between groups (MS vs Control) using independent t-tests.

3.3 Results

Table 33 details the demographic data; there were no significant differences between the two groups ($p > 0.05$). There was a significant difference between groups of the resting thickness of TrAb ($p=0.02$, 95% CI -1.34-0.13) with the MS group having thinner resting TrAb (mean scores). There were no other significant differences between groups. Tables 33 and 35 detail the summary results for the thickness of TrAb and IO (respectively) at rest and activation. There were no significant differences between the MS and Control groups for IO rest ($p=0.75$,95% CI -1.73 to 1.27), TrAb at activation ($p=0.79$, 95% CI-1.16 to 0.89), IO activation ($p=0.91$, 95% CI -1.56 to 1.4), TrAb percentage increase ($p=0.78$, 95% CI-1.85 to 33.45) and IO percentage increase ($p=0.90$, 95% CI -9.73 to 8.08).

Demographic data	**People with MS**	**Matched controls**
	(mean/ standard deviation /range)	**(mean/ standard deviation /range)**
Height/ cm	170.1+/-12.7 range: 154-193cm	167.6+/-10.7 range: 152-185cm
Weight/kg	74.9+/-20.4 range: 50-109kg	70.1+/-14.6 range: 53-98kg
Body Mass Index (BMI)	25.4+/-3.9 range: 19-33	24.4 +/-2.72 range: 21-32
Age/ years	54.5+/-10.7 range: 40-77	54.3+/-2.7.4 range: 35-77
Matched control demographic data is normal (KS test); No significant differences between groups with independent t-test ; Sample size: n= 17 pwMS+ 17 matched controls		

Table 33: Demographic data for participants and matched controls

	People with MS			Controls		
	Mean/mm	**SD**	**Range (min-max)**	**Mean/mm**	**SD**	**Range (min-max)**
TrAb at rest	2.9	0.85	1.9-5.3	3.7	0.88	2.3-5.0
TrAb on activation	4.0	1.89	2.0-9.9	4.0	0.85	2.2-5.3
Percentage increase	26.70%	32.1	-20.4-95.6%	10.9%	14.2	-5.1-50.7%
MS= Multiple Sclerosis, TrAb=Transversus Abdominis, IO= Internal Oblique, SD= standard deviation						

Table 34: Summary of results for Transversus Abdominis at rest, on activation and percentage increase for people with MS and matched controls

	People with MS			Controls		
	Mean/mm	**SD**	**Range Min-max**	**Mean/mm**	**SD**	**Range Min-max**
IO at rest	6.6	1.91	3.8-10.4	6.8	2.36	3.3-10.4
IO on activation	7.2	2.07	4.3-11.1	7.2	2.1	3.7-10.1
Percentage increase	9.8	12.1	-10.4-42.9%	10.6	13.4	-17.7-38.3%
MS= Multiple Sclerosis , IO= Internal Oblique, SD= standard deviation						

Table 35: Summary of results for Internal Oblique at rest, activation and percentage increase for people with MS and matched controls

3.4 Discussion

This small exploratory study suggests that MS might affect the thickness of the TrAb muscle as significant differences were demonstrated between the resting thicknesses of TrAb of people with MS compared with matched controls. There were no other significant differences between the two groups in either resting or activation thickness of TrAb or IO. This is the first study to measure the thickness of TrAb or IO with USI in people with MS; hence there is no published data to compare our results with. A comparison with published data of healthy people and LBP is documented in table 36 page 276. Of note, the thickness of TrAb and IO of people with MS are in the region of 1-2 mm thinner than healthy populations. Potential reasons for this are discussed after table 36.

Author	**Study sample and activation method** (data presented for healthy or matched control).	**TrAb rest** Mean/mm (+SD and range)	**IO rest** mean/mm (+SD and range)	**TrAb ASLR** mean/mm (+SD and range)	**IO ASLR** mean/mm (+SD and range)	**Mean % increase TrAb** (+SD and range)	**Mean % increase IO** (+SD and range)
Fox et al (unpublished data (2011) Reliability study.	ASLR in healthy people, mean age 35.8yrs, BMI not recorded.	**3.5** (±1.1)	**8.1** (± 2.8)	**4.3** (± 1.2)	**8.3** (± 2.4)		
MS vs matched control (MC) data from this clinical trial (2012)	People with MS compared with MC in ASLR Mean age 54.5 yrs, Mean BMI 25.4	MS **2.9** (± 0.8) MC **3.7** (±0.8)	MS **6.6** (± 1.2) MC **6.8** (±0.8)	MS **4.0** (± 1.9) MC **4.0** (±0.9)	MS **7.2** (± 2.1) MC **7.2** (± 0.8)	MS **26.7%** (±32.1) MC **10.9%** (± 14.2)	MS **9.8%** (± 14.2) MC **3.7%** (± 13.4)
Critchley & Coutts (2002)	Comparison of LBP with MC mean age 32 yrs, mean BMI 22	**5.1** (± 1.2 range 3.0-7.1)	**9.3** (± 4.0, range 2-24.5)	N/A as differing methodology	N/A as differing methodology		
Rankin et al (2006)	Healthy mean age 33 yrs, mean BMI 26.2	**4.5** (±1.3 range 1.9-7.1 (male) **3.6** (± 0.9range 1.8-5.4 female)	**11.8** (± 2.7 range 4.8-15.6 male) **8.5** (± 2.2 range 4.1-12.9 female)	NR	NR		
Mannion et al (2008)	Healthy males age 40.5 yrs, females age 42.1 yrs BMI not reported	**4.0** (± 1.0 range 3.6-4.5 male) **3.6** (±1.0 range 3.4-4.0 female)	**8.6** (±2.4 range 7.5-9.7 male) **6.7** (± 2.1 range 7.3-2.4 female)	NR	NR		

ASLR= active straight leg raise. MS= Multiple Sclerosis. ADIM= abdominal drawing in manoeuvre. HC= healthy control. Yrs = Years. BMI= Body Mass Index. LBP= low back pain. Healthy= people with no disease or pathology. SD= standard deviation. Activation data only included if methodology was ASLR not ADIM. NR = not reported.
Note: Transducer placement varies between studies from measurements taken at mid axillary to 2.5 cm anterior to mid axillary line which may account for difference in thickness.

Table 36: Comparison of thickness of TrAb and IO with published literature. Continued over leaf

Author	**Study sample and activation method** (data presented for healthy or matched control).	**TrAb rest** Mean/mm (+SD and range)	**IO rest** mean/mm (+SD and range)	**TrAb ASLR** mean/mm (+SD and range)	**IO ASLR** mean/mm (+SD and range)	**Mean % increase TrAb** (+SD and range)	**Mean % increase IO** (+SD and range)
Teyhen et al (2009)	Matched controls for LBP study ASLR mean age 36.7 yrs, BMI 27.2	**4.4** (±0.1) range not reported	**8.7** (±3.0) range not reported	NR	NR	**23.7%**	**11.2%**
Kordi et . (2011)	Healthy mean age 27.8 yrs, mean BMI =24.3	**2.5** (±1.0) range not reported	**7.3** (±1.7) range not reported	NR	NR		
Gill et al (2012)	Athletes, resting thickness mean age 19.8 yrs, mean BMI= 24.9	**4.5** (± 0.8, range 4.2-7.8)	**10.7** (± 2.1, range 9.9-11.5)	NR	NR		
Teyhen et al (2012)	ASLR healthy mean age 21 yrs, mean BMI=25	**3.9** (±0.09) male **3.3** (±0.09) female range NR	**10.4** (± 0.23) male **7.5** (±0.14) female range NR	**4.3** (± 0.12) male **3.6** (±0.11) female range NR	**15.5** **(**± 0.29) male **8.0** (± 0.1) female range NR	**9.27%** (male) **9.39** % (female)	**10.49%** (male) **6.16%** (female)

ASLR= active straight leg raise. MS= Multiple Sclerosis. ADIM= abdominal drawing in manoeuvre. HC= healthy control. Yrs = Years. BMI= Body Mass Index. LBP= low back pain. Healthy= people with no disease or pathology. SD= standard deviation. Activation data only included if methodology was ASLR not ADIM. NR = not reported.
Note: Transducer placement varies between studies from measurements taken at mid axillary to 2.5 cm anterior to mid axillary line which may account for difference in thickness.

Table 36 continued

Physical Fitness

A possible explanation for the differences in thickness of TrAb of people with MS with matched controls in our study is their differing physical fitness and activity levels. The MS group had EDSS levels ranging between 4.0- 6.5; so that at best they could walk a maximum of 500 metres without an aid whilst at worst they required constant bilateral assistance to walk 20 metres. In contrast, the healthy control group were typically active and involved in a variety of types of exercise such as dance classes, tai chi, yoga, running and hill walking (anecdotally reported to the researcher). Weight lifters have been reported to have significantly thicker resting TrAb than matched controls (p=0.01) (Sitilertpisan *et al* 2011). Referring to table 36 the mean thickness of TrAb at rest for athletes was 4.5mm (Gill *et al* 2012) compared to the MS sample which was 2.9mm. Participants measured by Teyhen et al (2012) reported TrAb at rest of 3.9mm in army recruits with a mean age of 21 years, further signifying that physical activity may be a contributing factor towards to the differences measured between the resting thickness of TrAb of people with MS and matched controls.

Reliability

There are several factors which may have contributed to inaccuracy of the results. Whilst the reliability study demonstrated intra-rater reliability was high, this was performed on a sample of young healthy physically active people; reliability may have differed for the MS sample. Since the reliability study was performed, new research by English et al (2012) has suggested that USI may not be reliable for measuring the cross sectional muscle area in people with neurological pathology due to an alteration in the muscle composition resulting

from sarcopenia. Deconditioned muscle can appear hyperechoic which can make it difficult to accurately measure thickness as defining the fascial borders of the muscle becomes more difficult. Parkkola et al (1993) reported that fat infiltration as a result of either disuse atrophy or pathology may make the muscle look thicker upon USI but not actually be a result of thickness changes. However their results contrasted with the findings of our study.

On reflection a reliability study conducted in people with MS prior to the images being taken may have increased the confidence in the measurements taken for comparison with matched controls.

Body Mass Index

Mannion et al (2008) state that BMI can cause the thickness of TrAb to increase. Anatomical sites with a tendency to fat deposition can affect reliability (English et al 2012). Obesity is also a limitation to the use of USI (Pretorius and Keating 2008). Our participants were matched in terms of age and size with no significant differences between BMI; furthermore images taken from participants with high levels of subcutaneous adipose tissue were excluded. However, BMI alone does not indicate deposition of adipose tissue. Images were excluded from the analysis if the thickness of adipose tissue prevented the measurement of the bottom (deepest) fascia of TrAb. In some cases the deepest layer was not visible and in some cases the presence of adipose within the muscle distorted the US image making it impossible to find a precise point to measure.

Symmetry

In this study an image was taken contralateral to the ASLR. In the case of the people with MS the ASLR was performed using the participants stronger leg and in matched controls the image was taken on the side to match. Other researchers have taken bilateral measurements of the deep abdominal muscles (Teyhen *et al* 2011), however evidence suggests that relative symmetry exists in the thickness of both TrAb and IO between the two sides of the trunk in a variety of populations including: healthy people (Rankin et al 2006; Mannion et al *2008)*, athletes with a tendency to be one sided such as rowers (Gill *et al* 2012) and cricketers (Hides *et al* 2008), and amputees (Springer and Gill 2007). Springer and Gill (2007) demonstrated that there were no significant differences between the resting thicknesses of TrAb between both sides of the trunk. Activation of the deep abdominal muscles has also been reported bilaterally in the presence of stroke (Marsden *et al* 2013), hence the decision to US one side contra-laterally to the ASLR in our study. However, it is possible that imaging bilaterally, as described by Teyhen et al 2011, may have further assisted in the understanding of the deep abdominal muscles.

Variability

In considering the thickness of the abdominal muscles, averages of measurements can disguise the variability of results. Variability within both samples (people with MS and controls) was evident (see figure 11 and figure 12). Greatest variability was demonstrated in the percentage activation increases; some individuals displayed negative percentage increases, meaning that the muscle appeared thinner with activation. When assessing the raw data with the supervisory team, any participants who displayed negative percentage

increases or measurements which lay outside of the mean score ± two standard deviations of the mean were re-measured by the researcher (EF) to ensure credibility of the results. Rankin et al (2006) similarly reported large variations in a sample of healthy people (n=123, see table 36 page 276), with resting TrAb and IO thickness reported between 0.9-7.1mm and 4.1-15.6mm respectively. Critchley and Coutts (2002) reported also wide variability in US thickness measurements of TrAb within their healthy control group. Rankin et al (2006) further reported that measurement error of 1-2mm can occur when measuring US scans of abdominal muscles. This seems a high level of measurement error when considering the thickness measurements of TrAb reported in the literature can be as low as 2.5mm (Kordi *et al* 2011) and 3.9 mm (Teyhen et al 2012).

A further potential contributing factor to the variability of the results within the MS sample is the differing EDSS levels of participants. Given the eligibility criteria of EDSS 4-6.5, the range in walking ability of participants was considerable (refer to page 265). As previously discussed activity levels may impact on the thickness of TrAb (Sitilertpisan *et al* 2011) hence it is plausible that the variability of the results could, at least in part, be attributed to differing activity levels.

Other factors to take into account when finding matched controls includes the parity of female participants. In this study female participants who were nulliparous were matched to participants who had given birth within the last five years. Critchley and Coutts (2002) reported that TrAb USI were thinner in woman who had given birth than those who had not.

The results for the thickness of TrAb and IO in this MS sample and matched controls were similar to published research (see table 36 page 276). One factor that might have accounted for the differences observed between MS and conrols may have been the transducer position. The protocol that was used located the transducer at midline between the iliac crest and lower rib as Teyhen et al (2011). Other researchers reported placing the transducer at 2.5cm anterior to mid axillary line to take images; this is visualising a different section of the muscle (Critchley & Coutts 2002; Kordi et al 2011).

Negative percentage increases

Negative percentage increases may have occurred due to compression of the muscle, either by operator error as a result of unduly pressing the transducer against the abdominal wall (Ishida & Watanabe, 2012) or as a result of visceral structures such as a full bladder compressing them muscles and making them appear narrower (Teyhen et al 2007). Alternatively some individuals may preferentially activate IO and/ or EO which could compress TrAb, making it appear narrower (Teyhen et al 2007). It is well documented that neuromuscular dysfunction can exist in TrAb as a result of LBP (Hodges & Richardson 1996; Hodges 1999; Ferreira et al 2004; Hides et al 2008), which may result in dominance of IO and /or EO in order to stabilise the spine (Silfies et al 2005;Brown & McGill 2010). Our participants were screened and excluded if they self-reported any episodes of LBP within the last three months, however Critchley and Coutts (2002) reported that neuromuscular dysfunctions can persist for two years after an episode of LBP, when the person is asymptomatic. Future studies assessing deep abdominal muscles could exclude potential

participants if they report episodes of LBP within the last two years. From a pragmatic perspective, recruiting a large enough sample size of matched controls of an appropriate age which have not had any episodes of back pain within the last two years however could be problematic.

Research in the field of LBP has used EMG to quantify the onset of activation of TrAb (Hodges & Richardson 1999), considering that it is the onset of muscle activation in response to the anticipation of movement which becomes dysfunctional (Vasseljen *et al* 2009). Our exploratory study did not quantify this phenomena and hence no conclusions can be drawn regarding the onset of activation on the deep abdominal muscles. Future research could potentially investigate the order of onset of TrAb, IO and EO, however, invasive fine wire EMG or high-resolution m-mode USI would be required (Mannion et al 2008).

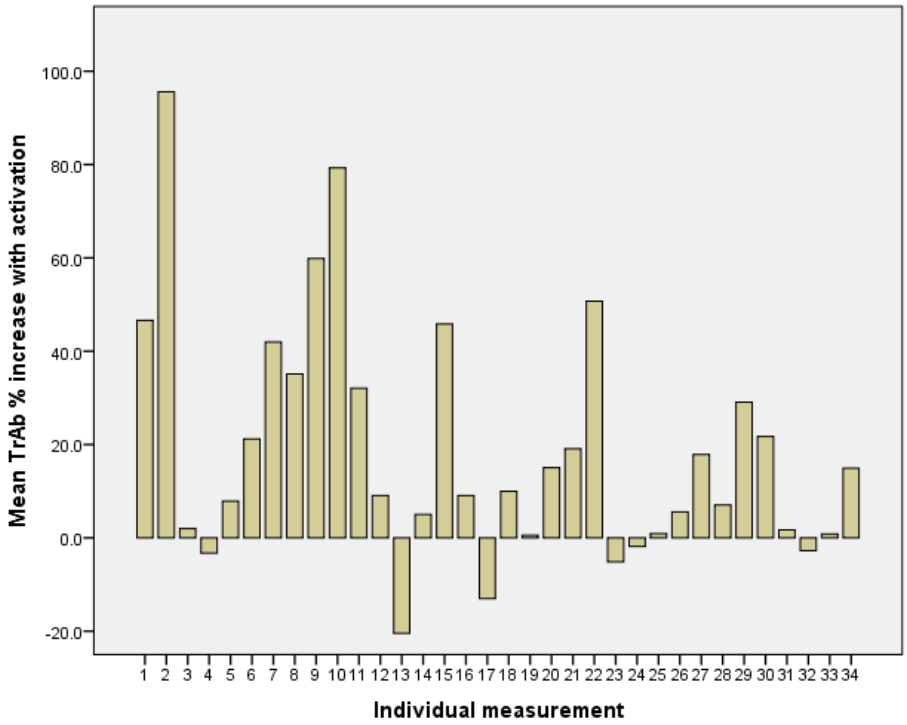


Figure 11: The variability of individual activation percentage increases in Transversus Abdominus (both groups included)

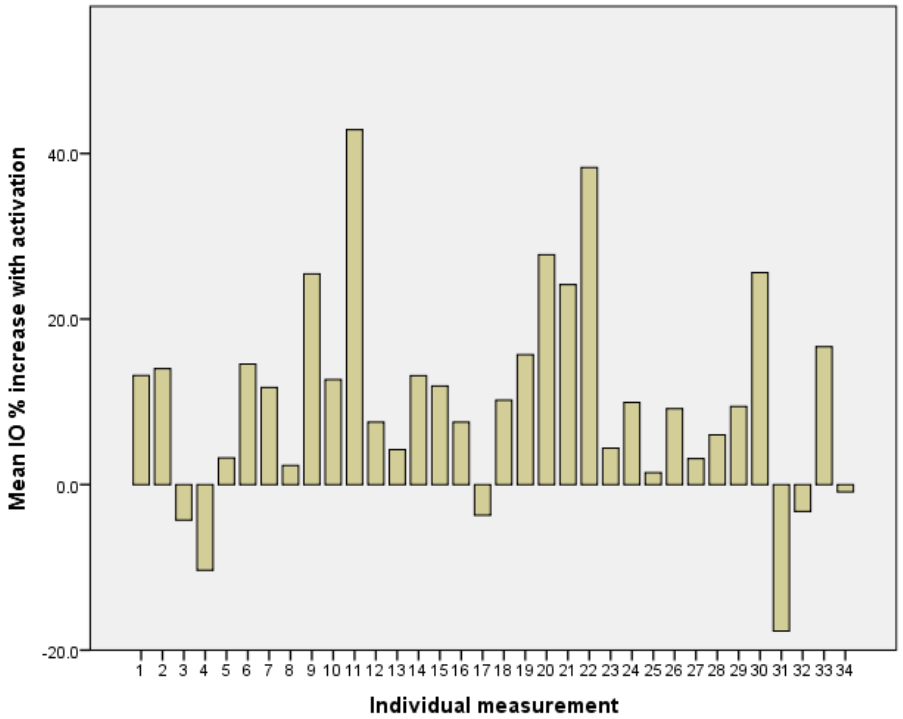


Figure 12: The variability of individual activation percentage increases in Internal Oblique (both groups included).

Age

Age may have affected the variability of the results. The age range of the sample was 35-77years. Panjabi (1992) proposed that the importance of the stabilising muscles decreases with age as the osseous and ligamentous spine becomes less flexible and therefore more stable. However evidence suggests that aging may not affect the deep abdominal muscle cross sectional area, with one study reporting that USI of TrAb at rest of active elderly people (mean age 85years) was not significantly thinner than their 20year old counterparts (p=0.01) (Ikezoe et al 2012). Further to this Rankin et al (2006) suggested that correlations between muscle size and age were too low to be of clinical significance (r<0.42). In a sample of older adults (n=12, mean age 75 years) resting thickness of TrAb upon USI was 5.6mm (±0.15) which was comparable to that reported in younger populations (see table 36 for data) (Stetts et al 2009).

In light of this the variability of the results seen in this sample is unlikely to be attributed to the range of ages.

3.5 Conclusions drawn from comparing the deep abdominal muscles of people with MS with matched controls

These results suggest that people with MS have thinner TrAb at rest than matched controls. There were no significant differences found in any of the other measures. Consistent with previous findings in other populations, there was substantial between-subject variability with respect to absolute muscle thickness at rest and percentage increase during activation.

The reliability of USI may be less in people with MS than in matched controls with possible impact upon the responsiveness of this measure in this patient group. Consequently conclusions on abdominal muscle activation should be drawn tentatively.

Section Two, Chapter Four: The effects of Pilates upon the deep abdominal muscles of people with MS

4.1. Introduction

A key component of the Pilates method is the specific training of the deep abdominal muscles in order to improve core stability (Dorado *et al* 2012). Whilst the effect of Pilates upon measures of function has been empirically evaluated (see chapter 3 page 66 for full literature review), few studies have been performed to establish an understanding of change at the level of impairment. The predominance of the existing research has been performed in healthy people (Cruz-Ferreira et al 2011: a systematic review) and in a clinical populations of people with LBP (Pereira et al 2011). Hence little is known about the effect of Pilates upon people with neurological conditions, despite the fact that core stability training is commonly advocated as physical therapy.

Questions have arisen regarding whether it is the magnitude of activation (Critchley *et al* 2011), hypertrophy (Dorado et al 2012) or onset of activation (Vasseljen et al 2012) (or indeed a combination of all three) of the deep abdominal muscles that is responsible for the improvements in function noted with Pilates and core stability training.

Research has demonstrated that a delay in activation of TrAb is associated with LBP but such research has not been performed in people with MS. Looking at a neurological population, in specific supratentorial stroke, Marsden et al (2013) reported that there were no significant differences ($p=0.08- 0.19$) in the magnitude of activation of TrAb or IO in response to a head lift whilst supine.

This was visible in both the ipsilateral and contralateral side to the stroke. As previously discussed, an explanation for this is that the trunk is bilaterally innervated (refer to page 270). This suggests that these muscles could be trained to improve trunk stability, as is the intention of the Pilates /core stability training interventions currently used by therapists in MS clinical practice (Freeman *et al* 2012). However the study by Marsden et al (2013) assessed magnitude of response with invasive EMG and not onset of activation or magnitude of activation with USI. It is non-invasive USI measurements that our has study focused on.

This chapter describes the exploratory study undertaken to investigate the effects of Pilates exercise on deep abdominal muscle thickness and activation measured by USI.

4.2 Literature review

This section will provide a summary of the available evidence pertaining to changes occurring in the deep abdominal muscles in response to Pilates and/or core stability exercise intervention. The literature reviewed will include data from healthy people and people with LBP due to the dearth of related research in any neurological condition.

Search strategy

In order to focus specifically on evidence pertaining to the effects of Pilates upon the deep abdominal muscles, the search strategy was as follows. The search engines 'Embase' which includes Ovid Medline and PsycArticles, CINAHL and 'Google Scholar' were searched from 1974 to 17th November 2014.

Using the terms as key words yielded the following results:

1. ‘Pilates’ and ‘deep abdominal muscles’ and ‘ultrasound’ or ‘sonography’,n = =1

2. ‘Core stability’ and ‘deep abdominal muscles’ and ‘ultrasound’ or ‘sonography’, n =3

3. ‘Pilates’ or ‘core stability training’ and ‘lateral abdominal wall’ and ‘ultrasound’, n = 1

4. ‘Pilates’ or ‘core stability training’ and ‘transversus abdominis’ and ‘ultrasound’, n = 6

5. ‘Pilates’ or ‘core stability training’ and ‘internal oblique’ and ‘ultrasound’, n = 1

6. ‘Pilates’ or ‘core stability training’ and ‘transversus abdominis’ and ‘change’, n =4

7. ‘Pilates’ or ‘core stability training’ and ‘internal oblique’ and ‘change’, n =0

8. ‘Effects of Pilates’, n = 45

The results were sorted by relevance to include papers which were published in peer reviewed journals, in English, and those which measured USI of TrAb and IO. As described previously (page 236), EO was not included. A manual search was also performed by reading the reference lists of key papers.

The immediate effect of core stability exercises upon the deep abdominal muscles

Pilates exercises influence the deep abdominal muscles (Herrington and Davies 2005) and thickness increases of TrAb and IO have been reported using USI (Endleman and Critchley 2008). Trunk strengthening exercises performed in

conjunction with voluntary activation of the deep abdominal muscles with ADIM (similar to those used in Pilates) demonstrated thickness increases upon USI in TrAb and IO in a sample of healthy people (n=120) (Teyhen *et al* 2008), indicating that combining ADIM with trunk exercises activates TrAb. In a further study of healthy people (n=26) TrAb and IO thickness were shown to increase during Pilates in comparison to rest (TrAb $p< 0.001$, IO $p<0.01$) (Endleman and Critchley 2008). It is noteworthy that when the exercises were performed, as described by Endleman & Critchely as 'incorrectly' (without the voluntary activation of the deep abdominal muscles), there were still significant differences between rest and exercise US images of TrAb and IO ($p=0.01$). There were no differences between TrAb and IO thickness when performed 'correctly' or 'incorrectly' ($p=0.117$). This suggests that perhaps the voluntary recruitment of the deep abdominal muscles is not required in order to activate them. These results were mirrored by a small study using fine wire EMG in the TrAb of healthy people (n=9); activity occurred in the TrAb during stabilisation exercises both with and without voluntary activation by ADIM, there was significantly ($p=0.042$) more activity in TrAb with instructions to hollow (Bjerkefors *et al* 2010). Both studies were performed in a small sample of healthy people on one occasion. Although this research informs us that Pilates exercises activate the deep abdominal muscles (whether performed correctly or incorrectly), it does not investigate the influence of Pilates training over time, or whether any changes seen are clinically significant.

The effects of core stability training upon the abdominal muscles: healthy people

Pilates training may result in hypertrophy of both the deep abdominal muscles and RA. In a sample of nine healthy females, a programme of 36 weeks of

Pilates resulted in increased resting volume upon MRI of RA (21% $p<0.05$) and combined volume of TrAb and IO (8% $p<0.05$ pre-post intervention) (Dorado *et al* 2012). The greater percentage increase in RA may be attributed to exercises involving repeated trunk flexion in Pilates. The limited methodological rigor of this study makes it difficult to draw definite conclusions however it suggests that 36 weeks of Pilates may induce hypertrophy of RA and the deep abdominal muscles. It is noteworthy that the authors state that they were measuring hypertrophy as opposed to magnitude or onset of activation and as measures were taken at rest, the influence of activation (whether increased or delayed) was not measured.

Activation of the abdominal muscles may be voluntary or automatic and it is not established whether specific Pilates exercises are required to generate changes. The effects of a programme of eight weeks of Pilates compared with general strength training was performed in healthy people (n=34) (Critchley *et al* 2011). USI measurements were taken pre and post training at rest and whilst performing Pilates. People assigned to the Pilates intervention had increased thickness of TrAb ($p=0.007$) during 'the hundreds' (a supine flexion based Pilates exercise to voluntarily activate TrAb and RA). However there were also other significant differences in TrAb or IO in either interventional group over time. The strength training group had thicker IO than the Pilates group post intervention suggesting that generalised strength training may require IO activation to stabilise the trunk. There were no differences between strength and Pilates training for TrAb post intervention, suggesting that generalised strength training may be sufficient to activate TrAb without the necessity for specific TrAb training with the use of voluntary recruitment by ADIM. Despite

being a small study, good methodological rigor was employed in concealment allocation of randomisation and blinding. The results suggest that in healthy people, Pilates may increase the thickness of the deep abdominal muscles but only during the performance of the exercises, leaving the superiority of Pilates over general strength training questionable. Furthermore no functional outcome measures were taken so the effect of Pilates upon function cannot be concluded.

The effects of core stability training upon the abdominal muscles in clinical populations: neurological conditions and low back pain

To date there has been no research published to evaluate the effect of Pilates or core stability exercises upon the US characteristics of TrAb and IO in people with any neurological condition, including MS. Hence literature taken from the field of musculoskeletal physiotherapy has been included here in order to assess the effects upon a clinical population.

Research to assess the effects of exercises upon TrAb and IO has been performed in people with back pain, using USI to measure thickness changes. A sample of 109 people with LBP were randomised to eight weeks of either general or 'sling' exercises (in the sling exercises the body weight was supported by slings and the abdominal muscles were voluntarily recruited with the ADIM). Results demonstrated that increases in TrAb contraction thickness were weakly correlated (R^2= 0.10) with reduced pain. The authors concluded that eight weeks of training using slings or general exercises generated only marginal changes in the contraction thickness of the deep abdominal muscles (Vasseljen and Fladmark 2010). Eight weeks however may not be sufficient

duration of time to generate hypertrophic changes of the muscles. For example, some research has demonstrated that eight to 12 weeks of intensive strength training is required in order to produce muscle hypertrophy in the paraspinal muscles (Danneels *et al* 2001). Hence any thickness changes occurring may be due to neural factors influencing increased activation.

In LBP it is considered that it is the onset, rather than the magnitude of activation or hypertrophic changes of the deep abdominal muscles, which are associated with pain (Hodges & Richardson 1999). A further publication suggests that core stability training may not influence the onset of activation (Vasseljen *et al* 2012). Using m-mode USI to measure onset of activation, abdominal muscle onset was shown to be 'largely unaffected' by eight weeks of core stability exercises in a sample of 109 people with LBP (presumably from the same sample as the aforementioned study, Vasseljen & Fladmark 2010) randomised to either core stability training, 'sling exercises' or general exercises; with no correlations to changes in pain. These studies suggest that a programme of core stability training may have little influence upon the deep abdominal muscles in people with LBP, or any clinical measures.

Generally, EMG has been used for measuring onset of activation of the deep abdominal muscles. More recently high-resolution m-mode USI has been shown to have preliminary validity as a method of assessing onset of activation (Mannion et al 2008). The validity was not ascertained at the time of designing hence it was not used in my study, but studies using this methodology may provide a useful non-invasive method to provide deeper understanding of deep abdominal muscle activation patterns.

To date the research evidence assessing the effect of exercise upon the deep abdominal muscles has been derived from small mechanistic studies. These have used EMG or USI in either healthy people or those with LBP. Whilst these studies contribute towards the understanding of the role of the core stabiliser muscles in response to exercise, little is understood of the behaviour of the deep abdominal muscles in people with neurological pathology.

Due to the paucity of published literature regarding the effects of exercise upon the deep abdominal muscles of people with MS, the aim of this exploratory study was to assess the effects of Pilates exercises compared with Standardised physiotherapy and Relaxation (placebo) exercises upon the thickness of TrAb and IO at rest, and activation during contralateral ASLR.

4.3 Methods

Recruitment and eligibility criteria

Recruitment and eligibility is described in methods for the clinical trial (page 144). USI was performed upon the first 22 consecutive people with MS recruited at the Plymouth centre. Demographic data regarding age, sex, height, weight, BMI and diagnostic data including years since onset was collected as per the main study.

Randomisation and blinding

Participants were allocated to intervention groups with concealment allocation (as per the clinical trial). USI was performed after written consent was taken. The procedure used to obtain and measure the US images is detailed on page

258. The USI was performed by the researcher (EF). There was no blinding of the assessor at the point of acquiring images. However to ensure blinding during measurement of the images, the image clips were given to the supervisor (AH) who deleted identifiable data and re-coded the clips such that it was not possible to know whether they had been taken pre or post intervention. The clips were then returned in a randomised order for the researcher to measure.

In the case of visually identified outliers, raw data (US scans) were re-measured. In the case of unclear images advice was sought from the academic supervisor (AH).

Normalisation of US cross- sectional thickness of the abdominal muscles according to body mass

In comparing US measurements of the cross sectional thickness of muscle between participants, body mass is a factor which requires consideration. To date the reporting of USI measurement of the abdominal muscles has tended to be either as absolute values (mm), or as a percentage change between rest and activation (Teyhen et al 2008, 2009)

Normalisation of data can change the outcomes of statistical tests and hence is important to consider when reporting data (Nuzzo and Mayer 2013). Normalised data has been reported by some authors; Rankin et al, (2006) normalised abdominal muscle data using ratio scaling, but only reported the absolute values. The effect of body mass upon abdominal muscle thickness has not been dismissed with researchers either correlating the BMI with abdominal

muscle size or assessing the effect of BMI on abdominal muscles by adding BMI as a co-variant to the analysis. BMI was positively associated with TrAb thickness at rest ($r=0.66$, $p<0.001$) (Springer et al 2006). Teyhen et al (2012) found that TrAb thickness was equivalent in men and women when height and weight were controlled for, although the absolute values were different, with males being thicker. This was mirrored by Rankin et al (2006) who reported that men had larger abdominal muscles than women, but TrAb was not affected by normalisation by body mass.

Normalisation of the data can be performed by ratio scaling (also termed isometric scaling) and is calculated by dividing the muscle size measurement (cross sectional thickness/mm) by the body mass/ kg. The reported normalised data for abdominal and lumbar mutifidis muscle US measurements were calculated using ratio scaling (Rankin et al 2006; Kiesel et al 2007). However a recent publication has argued that the use of ratio scaling is inappropriate for normalising TrAb and IO US data (Nuzzo and Mayer 2013), proposing allometric scaling to be a more appropriate method.

Allometric scaling is based on the theory of geometric symmetry, in which humans have basically the same shape but differ in size. Calculated by dividing the physiological measurement by the body mass raised to an exponential power (the allometric parameter), allometric scaling assumes a curvilinear relationship between the physiological measurement and body mass (Nuzzo and Mayer 2013). It is noteworthy however that the normalisation of TrAb thickness upon USI was not considered by Nuzzo to be necessary with allometric scaling. This may not be the case for IO however.

Statistical analyses

Demographic data was summarised using descriptive statistics. The mean thickness and standard deviations were calculated using Microsoft Excel (2012). Thickness changes were also expressed as a percentage, in the same manner as the matched control data, which were calculated from the equation (Teyhen *et al* 2009):

$$\text{Percentage change} = \left(\frac{\text{activation thickness - resting thickness}}{\text{resting thickness}}\right) \times 100$$

Statistical analyses were performed using IBM SPSS version 20 software. Significance level was set at p=0.05. Data was tested for normality using a Kolmogorov-Simonov test in SPSS.

A factorial 3 x3 repeated measures ANOVA was performed to determine the effect of exercise over time and between groups. Assumptions for ANOVA were met; the data was normally distributed and was interval data. Further analysis was performed by normalising the data using ratio scaling for the US muscle thickness measurements (Rankin *et al* 2006) using the following equation (Nuzzo and Mayer 2013):

$$\frac{\text{abdominal muscle thickness/mm}}{\text{body weight/kg}}$$

4.4 Results

The sample characteristics are described in table 37 There were no significant differences between groups for the demographic data ($p>0.05$). The US image clips of 17 people with MS were suitable for measurement at baseline. Three of the 17 participants were excluded during the trial due to ill health, relapse and commencing drug treatment hence it was not possible to obtain follow up US data. The descriptive measurements of the abdominal muscle thickness at the three time points are detailed in table 38 page 298.

	Pilates n=6		**Standard Exercise n=6**		**Relaxation n=5**	
Gender: n (%) female	5 (84%)		5 (84 %)		2 (40%)	
Type of MS: n (%)						
Relapse remitting	1 (17%)		2 (34%)		1 (20%)	
Primary progressive	1 (17%)		0 (0%)		1 (20%)	
Secondary Progressive	4 (66%)		4 (66%)		3 (60%)	
Age/ years: mean (sd)	56.00	(10.56)	55.67	(14.77)	58.00	(7.91)
Height/cm: mean (sd)	165.33	(5.85)	169.40	(12.28)	173.00	(15.91)
Weight/kg: mean (sd)	69.20	(13.91)	70.60	(24.37)	82.20	(24.18)
BMI: mean (sd)	25.00	(3.63)	24.20	(4.66)	26.80	(4.49)
Years since diagnosis: mean (sd)	12.83	(14.26)	24.00	(10.64)	6.60	(7.60)

Table 37: Sample characteristics for the ultrasound data

Factorial 3 x3 repeated measures ANOVA did not demonstrate any significant differences for TrAb or IO within group or between group changes (see table 39 for p values and effect size). Further analysis performed with ratio scaled normalised data did not produce any significant differences within or over time.

Baseline	TrAb rest	TrAb rest normalised	TrAb ASLR	TrAb ASLR normalised	% change	IO rest	IO rest normalised	IO ASLR	IO ASLR normalised	% change
All groups combined n=17 (mean)	3.09	0.05	4.06	0.06	32.16	5.75	0.08	5.94	0.08	3.24
Sd	1.14	0.02	2.06	0.02	56.66	1.85	0.03	2.25	0.03	15.41
Pilates n= 6 (mean)	2.92	0.04	3.78	0.05	27.72	5.08	0.07	4.95	0.07	-4.10
Sd	0.31	0.01	1.51	0.02	46.04	1.58	0.02	2.21	0.02	22.53
Standard exercise n=6 (mean)	3.60	0.06	5.03	0.07	50.64	5.52	0.08	6.08	0.09	10.16
Sd	1.84	0.02	2.48	0.02	57.33	2.36	0.03	2.66	0.03	2.73
Relaxation n=5 (mean)	2.70	0.04	3.22	0.04	15.29	6.82	0.09	7.20	0.10	3.85
Sd	0.45	0.01	1.99	0.03	72.02	1.21	0.04	1.12	0.04	11.18
week 12	**TrAb rest**	**TrAb rest normalised**	**TrAb ASLR**	**TrAb ASLR normalised**	**% change**	**IO rest**	**IO rest normalised**	**IO ASLR**	**IO ASLR normalised**	**% change**
Pilates n= 6 (mean)	2.97	0.04	3.62	0.05	20.87	6.20	0.09	6.97	0.10	13.20
Sd	0.54	0.01	0.98	0.01	19.98	1.12	0.02	1.15	0.02	11.32
Standard exercise n= 5 (mean)	4.10	0.05	5.18	0.07	22.63	5.88	0.08	6.70	0.10	13.41
Sd	1.87	0.01	2.99	0.02	15.30	1.96	0.01	2.48	0.01	9.90
Relaxation n=3 (mean)	3.87	0.05	4.17	0.05	12.72	7.20	0.09	7.47	0.10	-1.18
Sd	0.76	0.01	0.46	0.03	38.12	1.39	0.05	3.82	0.08	38.02
week 16	**TrAb rest**	**TrAb rest normalised**	**TrAb ASLR**	**TrAb ASLR normalise**	**%change**	**IO rest**	**IO rest normalised**	**IO ASLR**	**IO ASLR normalise**	**% change**
Pilates n= 5 (mean)	3.30	0.05	4.48	0.07	36.34	6.54	0.09	6.76	0.10	3.57
Sd	0.54	0.01	2.26	0.05	66.93	1.36	0.01	2.51	0.05	34.49
Standard exercise n= (mean)	3.14	0.05	3.82	0.06	24.29	5.98	0.08	6.44	0.10	5.96
Sd	1.35	0.00	1.81	0.01	25.48	1.55	0.01	2.58	0.02	21.44
Relaxation n= 5 (mean)	3.06	0.04	3.54	0.04	13.42	8.74	0.13	7.76	0.10	14.32
Sd	0.74	0.01	1.22	0.02	13.69	6.72	0.14	1.06	0.04	47.20

Measures of thickness of deep abdominal muscles on ultrasound of real time and normalised measures (real time = mm, normalised= mm/ bodyweight), ASLR= active straight leg raise, sd= standard deviation

Table 38: Ultrasound thickness measurements in mm of Transversus Abdominis (TrAb) and Internal Oblique (IO) at baseline, week 12 and week 16.

Factorial 3x3 repeated measures ANOVA						
Effect of intervention over time	TrAb rest		TrAb ASLR		TrAb% change	
	p value	effect size	p value	effect size	p value	effect size
Within subject (exercise group)	0.26	0.24	0.09	0.35	0.67	0.11
Between-exercise group comparisons	0.68	0.07	0.68	0.07	0.45	0.15
Effect of intervention over time	IO rest		IO ASLR		IO % change	
	p value	effect size	p value	effect size	p value	effect size
Within subject (exercise group)	0.21	0.27	0.11	0.36	0.52	0.17
Between-exercise group comparisons	0.25	0.24	0.42	0.18	0.66	0.09
Legend: TrAb: Transversus Abdominis, IO: Internal Oblique, ASLR: active straight leg raise						

Table 39: Mixed factorial 3X3 repeated measures ANOVA

Factorial 3x3 repeated measures ANOVA with normalised data				
Effect of intervention over time	TrAb rest normalised		TrAb ASLR normalised	
	p value	effect size	p value	effect size
Within subject (exercise group)	0.31	0.25	0.16	0.29
Between-exercise group comparisons	0.31	0.23	0.65	0.09
Effect of intervention over time	IO rest normalised		IO ASLR normalised	
	p value	effect size	p value	effect size
Within subject (exercise group)	0.28	0.26	0.20	0.33
Between-exercise group comparisons	0.38	0.19	0.42	0.20
Legend: TrAb: Transversus Abdominis, IO: Internal Oblique, ASLR: active straight leg raise				

Table 40: Mixed factorial 3X3 repeated measures ANOVA with normalised data

4.5 Discussion

Summary of results

To my knowledge, this is the first study to evaluate the effect of Pilates exercises upon the resting thickness and activation of the deep abdominal muscles of people with MS using USI. Repeated measures ANOVA did not yield any significant differences for either TrAb or IO for either within group or between group measures. Using normalised US data produced similar results. This was an exploratory study with a small sample size, hence there is a possibility a type II error could have occurred.

Comparison to other studies

TrAb did not appear to be influenced by either exercise intervention. Other published research assessing changes in the thickness of TrAb and IO in response to Pilates or core stability training yield similar results. Critchley et al (2011) reported that healthy participants (n=34) randomised to eight weeks of Pilates or strength training did not demonstrate significant changes in muscle thickness at rest or during functional postures ($p=0.05$). Vasseljen & Fladmark (2010) reported that there were no significant changes ($p>0.05$) in thickness of either muscle after eight weeks of core stability or sling exercises in a larger sample (n=109).

Explanation of findings: Innervation of the abdominal muscles

It is feasible that people with MS may have reduced activation in TrAb and IO. Whilst there has been no research published to demonstrate that atrophy or delayed onset occurs in the deep abdominal muscles of people with MS, investigations have been performed in people with cerebral strokes. These

studies demonstrate bilateral activation of TrAb occurring with both symmetrical (head lift) and asymmetrical (unilateral hip flexion) tasks (Marsden *et al* 2013). This may be attributed to the bilateral innervation of the trunk muscles from the motor cortex and brain stem (Murayama et al 2001; Tsao et al 2008). Impairments in RA and EO have been noted post stroke (Dickstein et al 2004; Pereira et al 2011) which may suggest that there are differences in the neural control of the deep abdominal and superficial trunk muscles (Marsden *et al* 2013).

The participants recruited in our clinical trial were of an EDSS level 4.0-6.5 (moderately disabled) and whilst we did not have access to MRI scans to determine the exact location of sclerotic lesions, it would seem plausible that the deep abdominal muscles of this sample may not be impaired.

The intensity of the exercises performed in both Pilates and SE were highly unlikely to be of sufficient intensity to generate hypertrophic changes. In people with LBP, ten weeks of stabilisation exercises did not result in increases in the cross sectional area of paraspinal muscles upon computerised tomography (CT); higher levels of intensity training were decreed necessary to develop muscle bulk visible upon CT at rest (Danneels *et al* 2001). However, the question remains unclear whether hypertrophic changes are required in order to increase cross sectional area (visible upon USI or CT), or whether neural factors such as increased neuromuscular recruitment is sufficient to increase magnitude of contraction as measured by cross sectional area. Furthermore in evaluating the effects of exercise upon the TrAb and IO, it needs to be established which is more clinically relevant; resting thickness or percentage

change at activation. Percentage change (activation ratio) would encompass the aforementioned neural factors.

In the Relaxation group participants performed a progressive relaxation exercise, in which participants lay supine and systematically contracted and relaxed all the muscle groups of the body including the abdominal muscles. This was performed in a supine position without loading the abdominal muscles or destabilising the spine. It is very unlikely that this would result in changes of the deep abdominal muscles as a much higher intensity of load is required (Danneels *et al* 2001), however there is a possibility that this may have occurred.

Variability within the sample

This study described only the magnitude of activation as measured by the cross sectional thickness (previously discussed). The onset of activation, in terms of specific patterns was not evaluated. There was much individual variability within the sample. Variability in the response of the abdominal muscles to activation has been reported in various studies using both USI and EMG (Vasseljen & Fladmark 2010; Mannion et al 2008; Morris et al 2013). Vasseljen et al (2010) reported that 82% of the variability in TrAb and IO thickness was not attributed to LBP. Variability in the cross sectional thickness of abdominal muscles can be partially accounted for by body mass. Mannion et al (2008) reported that BMI can account for 20-30% of the variance documented in US measurements of TrAb (percentage change between rest and ADIM).

In healthy people a natural variance in muscle recruitment patterns has been reported (Morris *et al* 2013) suggesting that individuals employ differing

neuromuscular strategies to facilitate movement (Hu et al 2012). In people with LBP, EO has been noted to act as a dominant stabiliser in response to dysfunctional TrAb recruitment (O'Sullivan 2000). In a study by Westad et al (2010) using M mode USI to measure onset of activation in people with LBP, IO was found to be the first abdominal muscle activated in response to rapid arm flexions (prior to TrAb). Interestingly, Westad et al reported that IO has deep and superficial regions and the deep regions were activated prior to superficial regions.

It is possible that differing abdominal muscle recruitment strategies and EO dominance could occur in people with MS giving rise to the large variation seen in the sample. It is worth noting that Kordi et al (2011) reported the US thickness of TrAb and IO decreased significantly after food consumption. This was not controlled for in this study and may potentially have contributed to the variability of measures.

Limitations of the study

The intention of this study was to collect exploratory data regarding the behaviour of the deep abdominal muscles to aid understanding of reduced trunk stability and whether Pilates or SE improve impairments, specifically muscle activation. Whilst reliability studies were undertaken in healthy people, demonstrating high reliability, some US image clips had to be discarded from the analysis, rendering a smaller than initially anticipated sample size. Obesity and pathology can result in difficulties in acquiring clear images (English *et al* 2012) and five of the image clips were discarded due to poor image quality. The small sample size makes it impossible to draw definite conclusions regarding

the effect of Pilates upon TrAb and IO since a type two error may have occurred.

This study assessed only the magnitude of activation as measured by changes in cross sectional thickness, not the onset of activation. At the time of writing there is no research published to ascertain that people with MS have delayed onset of activation of TrAb and IO. Delayed onset of activation is considered to be a neural adaptation of the deep abdominal muscles, as seen for example in the presence of LBP (Hodges & Richardson 1999); future research could be directed at investigating whether delayed onset of activation also occurs in MS and furthermore whether therapeutic exercise affects this. Fine wire EMG inserted into the deep abdominal muscles is currently considered the only valid method of measuring onset of activation, which is invasive and can be painful (Hu *et al* 2011) and hence was not undertaken in this study.

USI motion (m) mode and tissue velocity imaging has demonstrated potential validity as a measure of the onset of activation of TrAb and IO in LBP (Mannion et al 2008; Vasseljen et al 2009; Westad et al 2010). In future research non-invasive m mode USI could provide a viable, non-invasive option for collecting data regarding the onset of activation of the deep abdominal muscles.

Measurement error

USI is a reliable and valid measure of abdominal muscle activity (McMeeken *et al* 2004) and can detect low level changes in muscle activity, as low as 12% MVC for TrAb (Hodges *et al* 2003). Ultrasound measurements are considered valid for measuring magnitude of change in TrAb and IO. However, USI does

not discriminate well between moderate and strong contractions of TrAb and IO (Hodges et al 2003). Implications of this are that the strength of TrAb and IO contractions which may increase in response to Pilates training may not be captured by cross thickness USI measurements. Further to this, when assessing the MDC, measurement error must be considered. Variations of 0.1mm to 0.48mm (Gnat *et al* 2012) in USI of cross sectional muscle thickness may be attributed to measurement error (Rankin *et al* 2006). In future studies the calculation of SEM, MDC and MCID, based on the mode of USI and transducer used, would enable clearer conclusions to be drawn regarding the clinical relevance of changes in these muscles.

4.6 Conclusions

This small scale exploratory study used USI to assess the changes of thickness in the TrAb and IO muscles of people with MS in response to a 16 week period of Pilates or Standard Exercises compared to Relaxation (placebo control). The TrAb was not affected by either exercise programme. Due to the small sample size and large variability caution is required when drawing conclusions. Moreover, this study measured only magnitude of activation by the cross sectional thickness of the muscle upon USI. Future research could be directed at measuring the onset of activation of the deep abdominal muscles in people with MS. This could be performed using EMG or by developing protocols for using non-invasive m-mode USI.

Section Two, Chapter Five: The Functional Reach Test, correlations with Ultrasound Imaging

5.1 Introduction

The Functional Reach Test (FRT) is a measure of balance which is commonly used in the clinical setting due to its ease of administration and performance, and its low cost (the only equipment required is a metre rule) (Liao and Lin 2008). The psychometric properties of the FRT are detailed on page 122. The FRT was developed in 1990 to be a dynamic measure of postural control (Duncan *et al* 1990). The ability to reach further is considered to be an indicator of greater postural stability and consequently better balance (Jonsson *et al* 2003). Recruitment of the deep abdominal muscles is considered to contribute to trunk stability (Vera-Garcia et al 2007) and potentially influence postural stability. This could theoretically affect the performance of a FRT. This chapter will therefore focus specifically on the interaction between activity of the deep abdominal muscles and FRT performance.

5.2 Literature review

Search strategy

In order to focus specifically on the evidence pertaining to correlations between functional reach measures and the deep abdominal muscles, the search strategy used is detailed here. The search engines 'Embase' which includes Ovid Medline and PsycArticles, CINAL and 'Google Scholar' were searched from 1974 to 17th November 2014. Using the terms as key words yielded the following results:

1.‘Functional reach’ and ‘abdominal muscle’, n = 0

2.‘Functional reach’ and ‘Transversus bdominis’, n =1

3.‘Functional reach’ and ‘Internal Oblique’, n = 0

4.‘Functional reach’ and ‘ultrasound imaging’, n = 84

The results were sorted by relevance to include: papers which were published in peer reviewed journals, in English and measured USI of TrAb and IO. As previously stated, EO was not included. In addition a manual search was performed by reading the reference lists of key papers.

Activity in Transversus Abdominis during functional reach tasks

USI imaging has demonstrated that reaching forwards activates TrAb, in both healthy people and those with LBP (n=18 + 18 matched controls) in comparison to standing (p=0.001) (Nagar *et al* 2014). This study was carried out using a blinded assessor with m-mode USI; the participants held a 4.6kg hand weight whilst performing the reaching task. This study contributes to the understanding of abdominal muscle recruitment during a functional reaching task, however the participants had LBP so we are no better informed as to the effects of neurological pathology on the activity of TrAb during these tasks. Similarly this study does not assess whether core stability training affects activity of the deep abdominal muscles during functional reaching tasks.

Conversely McGalliard et al (2010) reported that postural instability does not necessarily affect TrAb when associated functional reaching. In this study McGalliard et al (2010) measured the thickness of TrAb upon USI in both standing and functional reaching, with and without an ADIM. Despite the small

sample size (n= 16) of healthy people, the US measurements were blinded with high reliability (ICC 0.82-0.95) suggesting rigorous methodology. Although the results were not statistically significant, it is could be questioned as to whether US is a responsive enough to detect the magnitude of change that might be expected for TrAb activity during a functional reaching task. EMG would be required to detect low levels of activity within the muscle (Merlo *et al* 2003). In addition to this, measurement error can be 0.1- 0.48mm when measuring TrAb (Gnat *et al* 2012). With reference to research performed by Teyhen et al (2012) comparing the thickness changes of supine rest to automatic activation with ASLR, the thickness differences were minimal with resting TrAb being 3.3mm (±0.09) and ASLR 3.6 mm (±0.11). Similar (< 1 mm) differences between rest and ASLR were noted in my own reliability study (see page 257).

The proposed theory that core stability training can improve FRT has been supported with low level evidence in healthy samples (Johnson *et al* 2007). McPherson & Watson (2013) reported that one session of supine TrAb training with a clinician using US bio-feedback increased action in TrAb. This was measured using USI, whilst performing ADIM during a standing forward functional reach, at a first assessment and then five months later in asymptomatic adults (n=10, p=0.001). The authors attributed this to 'motor learning', however there were several methodological flaws which preclude the generalisation of this data to my study; the lack of control group, small sample size, and the sample comprised of healthy participants. Furthermore the participants were instructed to 'draw in the abdominal muscles' whilst performing the functional reaching tests. The FRT was not used as an outcome measure *per se*, rather as a standing functional activity in which the abdominal

muscles could be measured. It is proposed that motor learning occurs in the training of TrAb (O'Sullivan 2000) and performing stabilisation exercises such as the ADIM to train TrAb may improve function. Research assessing the effects of core stability training upon the TrAb and IO using USI to measure the automatic recruitment of these muscles during functional activities, such as reaching has not been published (at the time of writing).

The Lateral Functional Reach Test (LRFT) is a measure of medio-lateral postural instability (Brauer *et al* 1999), although the forward and lateral FRT measure different planes of instability they are moderately strongly correlated in healthy people ($r=0.65$, $p<0.05$) (DeWaard and Bentrup 2002). Although the trunk muscles recruited during a LRFT differ from those required to forward reach (Örtengren and Andersson 1977), the specific contribution of TrAb or IO in comparison to forward reaching has not been established.

This chapter aimed to examine the correlations between FRT scores and the thickness measurements taken at rest and during an ASLR of people with MS at three time points; baseline, week 12 (immediately post intervention) and week 16 (follow-up) as per the clinical trial.

5.3 Methods

Participants were recruited and USI of the deep abdominal muscles performed as detailed on page 144 and 258. A FRT (forwards and lateral) was performed by a blinded assessor as per the clinical trial protocol (described on page 102) at baseline, week 12 and week 16.

Analyses

Statistical analyses were performed using IBM SPSS version 20 software. Demographic data was summarised using descriptive statistics. The data was checked for normal distribution using Kolmogorov-Simonov. A Pearson's moment bivariate correlation was performed with two tailed significance set at p=0.05. Correlations were performed upon absolute values and data normalised with ratio scaling. Functional reach data was normalised (as suggested by Maranesi et al 2014) using the following equation (Hageman *et al* 1995):

$$\frac{\text{Functional reach score/cm}}{\text{Height/cm}}$$

The criteria for determining the magnitude of correlations was according to Cohen (1988) wherein >0.20 are weak, >0.50 are moderate, and >0.80 are strong correlations.

5.4 Results

Demographic, diagnostic and descriptive data are reported in table 37, page 297. The mean (±sd) for US thickness measurements is reported in table 38 page 299. The mean (± sd) for Functional Reach Tests is reported in table 41, page 313.

Correlations between normalised Forward Functional Reach and USI of TrAb and IO

At baseline there were weak correlations between TrAb during ASLR (r= -0.20, p=0.46) and normalised FFRT. At week 12 there were weak correlations between TrAb at rest (r=-0.36, p=0.22) and normalised FFRT. At week 16 there

were weak correlations between TrAb rest (r=-0.23, p=0.44) and IO rest (r=0.24, 0.41) and normalised FFRT. None of these were significant (p>0.05).

Correlations between normalised Lateral Functional Reach and Ultrasound imaging of Transversus Abdominis and Internal Oblique

At baseline there were weak correlations between IO at rest (r=0.36, p=0.17), IO during ASLR (r=0.27, p=0.32) and LFRT. At week 12 there were weak correlations between TrAb during ASLR (r=0.21, p=0.49), IO rest (r=0.23, p=0.45) and LFRT. There was a moderate correlation for IO during ASLR (r=0.54, p=0.06) and LFRT. None of which were significant (p>0.05). At week 16 there were weak correlations between TrAb rest (r=-0.29, p=0.31), IO rest (r=0.36, p=0.21), IO during ASLR and LFRT; none of which were significant (p>0.05).

Correlations between normalised Forward Functional Reach and normalised USI of TrAb and IO

Normalising the US data affected the results. At week 12 there was a weak correlation between IO rest (r=0.26, p=0.39) and FFRT. At week 16 there was a weak correlation between IO rest (r=0.39, p=0.16), and a moderate correlation for TrAb at rest (r=0.65, p=0.01) and during ASLR (r=0.55, p=0.04) and IO during ASLR (r=0.52, p=0.06) and FFRT. These correlations were significant for TrAb at rest (p=0.01)and ASLR (p=0.04) and FFRT.

Correlations between normalised Lateral Functional Reach and normalised USI of TrAb and IO

For the normalised USI and normalised LFRT data at baseline there were weak non-significant correlations for IO during ASLR (r=0.33, p=0.23) and a moderate correlation for IO at rest (r=0.50, p=0.05). At week 12 there were weak non-significant correlations between TrAb rest and LFRT (r=0.42, p=0.15). There were moderate correlations for TrAb during ASLR (r=0.62, p=0.02), IO at rest (r=0.61, p=0.03) and IO on ASLR (r=0.74, p=0.001) and LFRT; all of which were significant. At week 16 there were weak correlations between TrAb rest (r=0.45, p=0.10), TrAb ASLR (r=0.32, p=0.27), IO rest (r=0.48, p=0.08) and IO ASLR (r=0.30,p=0.30) and LRFT; all of which were non-significant.

Baseline	N	Minimum	Maximum	Mean	Sd
Forward functional each test (cm)	17.00	12.30	34.60	25.88	6.45
Lateral functional reach test (cm)	17.00	9.00	25.30	17.74	4.68
week 12					
Forward functional reach test (cm)	16.00	16.00	35.60	28.36	6.25
Lateral functional reach test (cm)	16.00	10.30	33.00	20.06	5.40
week 16					
Forward functional reach test (cm)	16.00	16.30	34.30	26.47	5.19
Lateral functional reach test (cm)	16.00	12.00	29.30	18.49	4.10
Sd= standard deviation					

Table 41: Functional Reach Test scores (mean and standard deviations of raw, non-normalised scores) of people who had Ultrasound Imaging.

5.5 Discussion

To my knowledge, this is the first study to correlate functional reach scores with US thickness measurements of TrAb and IO. There were increases seen in the mean functional reach scores for the whole group. Correlations between both deep abdominal muscle thickness and FFRT and LFRT demonstrated that the relationship between the distance reached outside of the base of support and deep abdominal muscle thickness ranges was generally weak. Potential explanations for these results are hereon discussed.

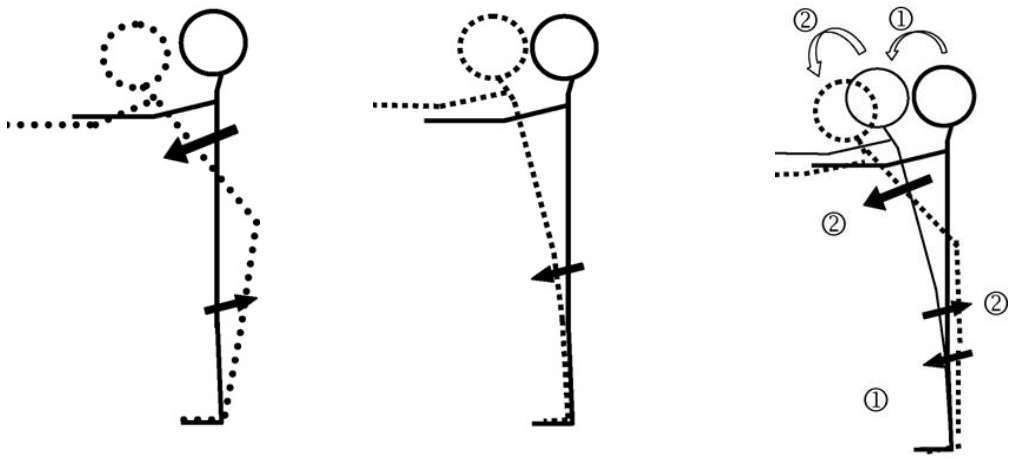
Movement strategies

Different movement strategies are performed by individuals. Forward trunk movements are accompanied by hip flexion and/or ankle dorsi-flexion (figure 13) and pelvic rotation (Liao & Lin 2008; Maranesi et al 2014). Studies show that individuals use differing movement strategies on different occasions, which may in turn affect the recruitment of the deep abdominal muscles (Morris *et al* 2013). It is reasonable to propose that activation of deep abdominal muscles may differ between those who reach using hip flexion or rotation strategies compared to ankle dorsi flexion strategies; future research could explore this.

Further to this people who are affected by diabetic neuropathy adopt different movement strategies in comparison to healthy controls (Maranesi *et al* 2014) suggesting that peripheral neuropathy can affect functional reach ability. Whilst there is no published data to draw upon, it is plausible that people with MS, who experience sensory impairments and weakness around the ankle joint (DeLuca *et al* 2004), may adopt a hip flexion strategy (thereby potentially activating the deep abdominal muscles) during a FRT. Maranesi et al (2014)

suggest that EMG evaluation of the muscles responsible for the differing strategies is needed to aid understanding of the anticipatory postural adjustments involved in the FRT.

This variability in movement strategies used by individuals in the FRT may, at least in part, help to explain the generally weak correlations seen in the results.



A B C

Figure 13 :Movement strategies for Forward Functional Reach Test.

A= hip flexion strategy, B= ankle dorsi flexion strategy and C = mixed strategy (Liao and Lin, 2008).

Balance is a result of a complex interaction of sensory and motor components (Tyson & Connell 2009). Contributions from the visual, proprioceptive and vestibular systems all contribute to maintenance of balance, and whilst the abdominal muscles may assist in stability they cannot be entirely accountable for changes in balance. Participants may have increased functional reach distances due to improvements gained in leg muscle power (Yahia *et al* 2011), improved sensory and motor strategies (Cattaneo et al 2007), reduced postural sway (Kaji et al 2010) and other factors not measured in this study.

In order to achieve medio-lateral postural stability, the contributions of the abdominal muscle groups may differ in comparison to the forward function reach. This has not been documented in the published literature but it could account for differences in lateral and forward reach correlations with abdominal muscle thickness.

Limitations

As discussed in the previous chapter, this study measured only magnitude of activation and not onset of activation; moreover the sample size was small for this exploratory USI component. Postural stability could be affected by the onset of activation and there is currently no published data exploring this in people with MS.

5.6 Conclusions

Deep abdominal muscle activation measured by thickness changes on USI in people with MS demonstrated generally weak, non significant correlations with forward and lateral reach scores. Normalising the abdominal muscle data with ratio scaling (normalising to body mass) resulted in significant correlations of moderate magnitude indicating that increased thickness of deep abdominal muscles is associated with increased reach distance. It is plausible that individuals use differing strategies in the performance of a FRT which could affect recruitment of the deep abdominal muscles. The literature to date regarding the activation of deep abdominal muscles when reaching is scant.. Further adequately powered studies are merited to determine whether this activation of the abdominal muscles affect reach distance.

Section Two, Chapter Six: Summary of USI of the deep abdominal muscles: combined literature and research findings

Reliability

USI is a feasible and relatively inexpensive modality for imaging the deep abdominal muscles. It has been extensively tested for reliability and has been found to be highly reliable when sufficient operator training is provided (Ferreira *et al* 2011). In excess of 20 hours of training is deemed to be the minimum required (Teyhen *et al* 2011). Factors affecting reliability such as transducer movement and pressure (Dupont *et al* 2001), position of participant (Arab *et al* 2013), method of activating the abdominal muscles (Brown and McGill 2010) and breathing mechanics (Ishida *et al* 2012) can all be controlled for by appropriate operator training. The majority of measurement error occurs when acquiring images rather than off screen measurement, further highlighting the need for operator training (Gnat *et al* 2012). In our study, the reliability of the researcher (EF) to acquire and measure images in healthy people was assessed prior to undertaking the experimental work. This was found to be high for measuring TrAb and IO at rest and during automatic activation with ASLR.

The effect of neurological conditions upon the deep abdominal muscles is not well understood, with very little research performed in people with MS. The exploratory USI was conducted as part of the clinical trial in order to gain some understanding of the effects of exercises upon the deep abdominal muscles. With hindsight it would have been preferable to undertake the reliability study (which informed the protocol development) on people with MS, rather than healthy controls. Measuring the onset of activation rather than just the

magnitude of change may also have enabled a deeper understanding regarding the underlying impairment and mechanisms of change in this patient group.

Validity

The validity of USI for measuring deep abdominal muscle activity has been ascertained (McMeeken *et al* 2004). The validity of measuring TrAb is confirmed for both voluntary and automatic activation, however for IO, validity appears to depend on using automatic activation strategies. Using USI to measure EO appears neither reliable nor valid due to the muscle geometry during contraction (John and Beith 2007). Both pathology and obesity can affect the validity of USI (Perkin *et al* 2003).

Responsiveness

The responsiveness of USI to measure clinically significant changes of the deep abdominal muscles has not been well established. Some authors report the SEM but the MCID has not been established for USI of the deep abdominal muscles.

Comparing the deep abdominal muscles of people with MS with matched controls

The results of our study demonstrated that TrAb was thinner in people with MS at rest, but there was no other difference between the resting or activation thickness of TrAb or IO of people with MS compared with matched controls. One potential explanation is the reduced general activity in this sample (EDSS 4.0-6.5). This concurs with research comparing physically active populations with matched controls (Sitilertpisan *et al* 2011).

The effects of Pilates upon the deep abdominal muscles

The TrAb and IO of people with MS does not appear to be affected by either Pilates or Standardised Exercises compared to Relaxation (control), which is in line with research in other conditions (Vasseljen and Fladmark 2010; Critchley *et al* 2011). The results however must be interpreted with caution due to the large variability noted and small sample size, which could potentially give rise to a type two error.

Correlations with Functional Reach Tests

There were weak to moderate correlations between the Functional Reach Tests and TrAb and IO. There is little literature which explores the influence of Pilates and/or core stability training upon deep abdominal muscles and how they may affect functional reaching. This finding is not unexpected since retraining balance when reaching requires a complex interaction of motor, sensory and proprioceptive control (Tyson & Connell 2009); thus it is influenced only in part by deep abdominal muscle activation.

Contributions to knowledge

This was the first study which explored the effect of exercise upon the deep abdominal muscles of people with MS using USI. Conclusions drawn from this research suggest that neither Pilates nor SE affect the magnitude of activation of TrAb or IO when assessed with automatic activation strategies. However, small sample sizes preclude generalising these findings due to the potential of a type two error.

Section Two, Chapter Seven: Future Research

7.1 Reliability of USI

The reliability of USI of the deep abdominal muscles has been well established in samples of healthy people (Koppenhaver *et al* 2009), however pathology can affect reliability (Perkin *et al* 2003). Reliability of USI has not yet been investigated in MS. Prior to any further research being performed using USI, reliability should be established in the target population for the protocol selected.

7.2 Development of protocols to validate the use of non-invasive imaging

Future research is needed to explore whether there are alterations in the pattern of activation of the deep abdominal muscles (for example delayed onset of activation) in people with MS. It is plausible to suggest that this may be the case. Currently methods of assessing onset of activation of deep muscles have been performed using invasive EMG (Vasseljen *et al* 2009), which can be painful (Hu *et al* 2011). In samples of people with LBP, high-resolution M- mode US has shown promising validity in measuring the onset of activation of deep abdominal muscles. Future research refining and validating M-mode and/ or Doppler US to image the activation patterns of the deep abdominal muscles would be justified to gain information regarding the behaviour and furthermore the effects of interventions.

7.3 Responsiveness

Another key area for future research is determining the clinical significance of such changes measured with USI. There has been very little research to define

this, nor to determine the SEM and MCID of US measurements in any populations. Furthermore, it is not known whether these changes would relate to hypertrophic changes at rest, at automatic activation or during voluntary contraction. Similarly it is not established whether neural factors such as increased recruitment or differences in onset of activation would result in clinically significant changes. Without such information the clinical application of research findings is problematic.

7.4 Functional measures

After developing protocols to establish the reliability, validity and responsiveness of both B and M-mode US in the target population, clinical research using larger samples could be implemented using a combination of functional outcomes measures and imaging to aid understanding of changes at the level of impairment. Such research could include assessment of other stabilising trunk muscles such as multifidus and external oblique.

Developing a deeper understanding of the responsiveness of USI is particularly warranted given the widespread implementation of therapeutic exercises targeted at improving trunk stability.

Section Two, Chapter Eight: Overall Conclusions of the Thesis

This chapter will focus upon amalgamating the findings of the clinical trial and ultrasound study to detail the contribution that this thesis has provided to the theory, practice and methodology, for the use of Pilates as an intervention to improve balance and mobility for people affected with MS. A summary of the results for the clinical trial is provided on page 225-6 and for the ultrasound study on page 319-320.

This was the first methodologically rigorous study to compare the effects of Pilates not only with a placebo (relaxation) but with an alternative form of exercise (Standardised Exercise {SE}). Furthermore, this was the first study to explore the underlying mechanisms of change using USI in a sample of people with MS undertaking exercise interventions. The clinical trial was designed to assess the use of Pilates as a method of core stability training, which has been widely implemented in clinical practice. Teaching voluntary activation of the deep abdominal muscles has been adopted by therapists, with the intention of improving balance and consequent mobility. The results of this thesis demonstrate that this specific approach is not required for improving balance and mobility in people affected by MS. As demonstrated by between group comparisons and effect sizes, both balance and mobility improved across a broader range of measures, and with a greater magnitude, in the SE group than the Pilates group. Whilst this does not negate the importance of abdominal muscle activity, it highlights that voluntary control of these muscles using the abdominal drawing in manoeuvre may not be necessary. The Pilates method employed in this clinical trial placed a heavy focus upon teaching the

participants to voluntarily activate the deep abdominal muscles. Therapists teaching the exercises to the SE group were expressly advised not to teach participants to voluntarily activate the deep abdominal muscles, hence abdominal muscle activity can be assumed to have been automatic.

In support of this the exploratory USI study demonstrated that during an active straight leg raise, the deep abdominal muscles activate in order to stabilise the spine, congruent with literature in this field (Hu et al 2011; 2012). This further indicates that teaching voluntary activation of the deep abdominal muscles may be redundant in improving balance and mobility as normal movement is sufficient in automatically activating the spinal stabilising musculature.

The results of this thesis can be implemented in both clinical practice and future research. In clinical practice, Pilates and core stability training could still be used by therapists and people with MS. However, focus may be better placed on the performance of task specific exercises and functional strengthening, in line with the exercises used in the SE intervention. With reference to the dose of exercise, the frequency (30 minutes once a week and 15 minutes of daily exercise) was sufficient to generate clinically and statistically significant changes beyond measurement error in walking speed and walking impairment as measured by the10 metre timed walk and the MSWS-12. Additionally statistically significant differences were achieved in balance as measured by the Functional Reach Test and Activities Balance Confidence Scale. Participants in both the SE and Pilates group were adherent to home exercise and there were no reported harms as a consequence of exercise. This indicates that these

forms of exercise can be safely recommended by therapists to people affected by MS.

With regard to the USI, prior to this thesis, no research had been performed using USI to measure activity in the deep abdominal muscles in people affected by neurological pathology. The research performed for this trial may contribute towards future protocol development for USI of the deep abdominal muscles and consequently improve the mechanistic understanding of disease upon activity in these muscles.

Appendix 1: Advert for SWIMS newsletter

RESEARCH WITH PLYMOUTH UNIVERSITY

Clinical trial for people with balance and mobility difficulties

Dr Jenny Freeman and Esther Fox, at the School of Health Professions, Plymouth University, are currently running a multi- centre clinical trial investigating different types of physiotherapy exercise for people with mobility and balance difficulties. They are looking for people who are experiencing mild to moderate difficulties with balance and mobility, and who are not currently participating in another clinical trial, to take part. If you are aged over 18 years, are able to walk independently with or without a walking aid such as a stick, and have not had a relapse within the past three months you may be eligible to participate in this study. The study will require your involvement for 16 weeks in total. You will participate in one to one exercise sessions with a physiotherapist over a 12 week period, during which time you will also be asked to undertake a home exercise programme. Over this time you will also be required to undergo three assessment sessions, involving tests of your balance and mobility. Your travel expenses for attending these sessions will be reimbursed.

If you would like further information please feel free to call Esther Fox, on

01752 587599

or email esther.fox@plymouth.ac.uk

Appendix 2: Invitation to participants

SWIMS Project Coordinating O
Clinical Neurology Research G
Room N7, ITTC Building 1
Tamar Science Park
PLYMOUTH
PL6 8BX.
Tel: 0800 015 3430 (FREEPHO

«Title» «Forename» «Surname»
«Address_1»
«Address_2»
«Address_3»
«Town»
«Postcode»

June 2011

Dear «Title» «Surname»

Re: Improving balance and mobility in people with Multiple Sclerosis

I am writing to let you know about a new research study that is being undertaken by Esther Fox and supervised by Dr Jenny Freeman who are both based at the School of Health Professions, University of Plymouth.

This letter is being sent to everyone who is registered on the South West Impact of Multiple Sclerosis (SWIMS) Project who is able to walk a short distance.

I would like to invite you to take part in this new study, which aiming to identify whether physiotherapy has an effect on balance and mobility. The enclosed information sheet explains the aims of this study. I would be very grateful if you could read this information along with the other enclosed documents. If anything is unclear, or you have any questions about the study, please feel free to ring Esther Fox on **01752 587599** or email her at esther.fox@plymouth.ac.uk to discuss your queries.

Thank you for taking the time to consider contributing to this study.

Yours sincerely

Prof John Zajicek (Consultant Neurologist, Chief Investigator for SWIMS Project)

Appendix 3 PARTICIPANT INFORMATION SHEET

Faculty of Health & Social Work
University of Plymouth
Peninsula Allied Health Centre
Derriford Road
Plymouth
Devon, PL6 8BH
United Kingdom

tel +44 (0) 1752 588 800
fax +44 (0) 1752 588 874
www.plymouth.ac.uk/healtheducation

PARTICIPANT INFORMATION SHEET

IMPROVING BALANCE AND MOBILITY IN PEOPLE WITH MULTIPLE SCLEROSIS (MS): A MULTI-CENTRE RANDOMISED CONTROLLED TRIAL.

Invitation to participate

We would like to invite you to participate in a new research study. Before you decide whether or not to participate, it is important for you to understand why the research is being done and what it will involve. This information sheet explains the background and aims of the study. Please take time to read it carefully and discuss it with family and friends or your own doctor or physiotherapist if you wish. If there is anything that is unclear, or if you would like more information, please ask us. Your participation in this study is entirely voluntary.

Why have I been invited?

You have been chosen because you are currently experiencing balance and mobility difficulties as a consequence of having MS. In total 100 people with MS, from 5 different centres throughout England and Scotland, will be participating in this research study.

What is the overall aim of the study?

Difficulties with balance and mobility are common in people with MS. These difficulties are wide ranging and may include unsteadiness when walking, standing, or undertaking tasks such as carrying a cup of tea. Physiotherapists use different exercise approaches when trying to improve people's balance and mobility. Currently we do not know which of these approaches is most effective in improving balance and mobility in people with MS. The aim of this study is to determine which of these three different exercise approaches is most effective in people with MS.

What will happen to me if I take part? What do I have to do?

If you choose to take part in this study your participation will be required for a total of 16 weeks.

In the first instance you will be asked to attend an assessment session. At this session, having had an opportunity to ask questions, you will be asked to

complete a written consent form. After doing so, you will be randomly allocated to one of three different physiotherapy interventions.

At this first session a physiotherapy team member, who is unaware of which exercise group you have been allocated to, will undertake an assessment of your balance and mobility. This assessment session will last for approximately thirty minutes. It will include measurements of the length of time you take to walk ten metres indoor; and how far you can reach forwards and sideways in standing. You will also be asked to complete 3 short questionnaires asking you about your mobility and balance. All of these assessments are commonly used by physiotherapists in their daily clinical practice.

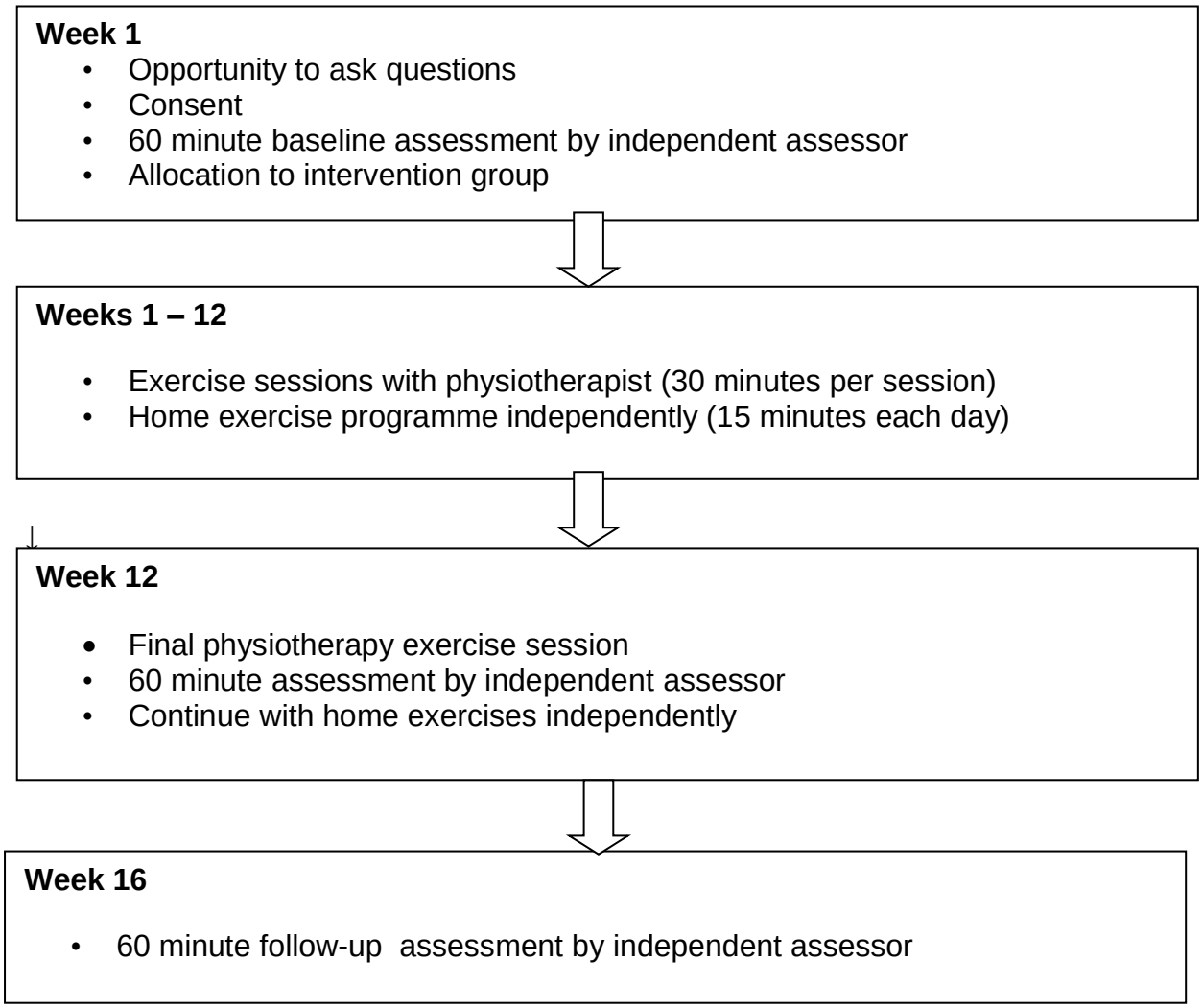
After undertaking these tests of balance and mobility we would then like to measure how effectively your abdominal muscles are working. No special preparation is required for this, although you are asked to wear comfortable loose clothing such as a tracksuit, so that your top can be rolled up in order for us to clearly see your abdominals. To measure the muscle activity we will ask you to undertake some movements of the arm, firstly when you are lying down and then when you are sitting up. During these movements we will use ultrasound scanning to measure your muscle activity. The measurements gathered using the ultrasound scanning will require that a small amount of gel is placed on the skin of the abdominal muscles where the ultrasound transducer will be placed. You should feel no discomfort whatsoever during this procedure. You will only be aware of movement of the transducer over the skin and a sensation of cold from the gel on your skin

Having undertaken these assessments you will then be given an appointment with the neurological physiotherapist who will undertake a programme of exercises with you over the next 12 weeks. The number of face to face sessions you will receive from your physiotherapist will depend upon the group to which you have been allocated. At a maximum you will be required to attend 12 weekly sessions with your physiotherapist, and at a minimum you will be required to attend three sessions. Each of these sessions will last for approximately 30 minutes.

Regardless of the group allocation, you will also be asked to undertake a brief home exercise programme (approximately 15 minutes) on a daily basis between physiotherapy sessions. A workbook will be provided to describe the exercises we would like you to practice, and you will be asked to complete a “tick-box” diary to record when you have undertaken these exercises. At week 12, the same balance and mobility assessments as undertaken at the beginning of the programme (the baseline assessment) will be repeated by the same assessor.

Four weeks after having completed the 12 week exercise programme (week 16) you will be asked to attend a final assessment session so that “follow-up” assessment can be undertaken. Once again, these assessments will be identical to those you completed at the first and 12th week of the study. As usual this assessment session will take approximately 30 minutes and will be undertaken by the same assessor.

A flowchart of this process is outlined below:



Will any expenses be paid?
Your travel expenses will be paid for your return journey to attend the physiotherapy treatment sessions and the three assessment sessions. Travel expenses will be reimbursed at a mileage rate of 40 pence per mile. Alternatively, if you are unable or unfit to drive you will have taxi fares reimbursed to attend these physiotherapy assessment sessions. The researchers will make and pay for the telephone calls to arrange your appointment.

Do I have to take part?
No. Participation in this study is entirely voluntary and if you decide not to take part your usual medical and physiotherapy care will not be affected in any way.

Will the study involve taking any new medication?
No. Other than the exercises, we will not change your existing medication or prescribe any new physiotherapy interventions during the 16 week period that you are involved in this study. You should continue to take all your usual medicines as prescribed; and to participate in your usual activities and exercise programmes.

Will I have to make any extra visits to my neurologist or GP?
You will not have to make any extra visits to your GP or to your neurologist. The only extra appointments you will need to make are to the physiotherapist, as described above.

If you decide to take part we will inform your GP by letter, with your permission.

What happens when the research study stops?
At the end of the study you will continue to receive the usual treatment that the physiotherapist provides to you.

Will my records be confidential?
All information collected about you during the project will be kept strictly confidential. You will be one of 100 people with MS that are involved in this project. You will be allocated a project number which we will use on all assessment records rather than your name or other identifying details. All information that we collect on you will be stored electronically on a computer which is password protected, in a document file that is also password protected. Your name and address will be stored separately from the other information you supply during the project so that you cannot be identified from your study records. If you choose to discontinue being involved in the study we will need to use the data you have provided so far so that we can analyse the results from the trial accurately. All information will be handled in compliance with the Data Protection Act (1998).

What are the potential benefits of taking part in this study?
By allowing these assessments before, during and after you have undertaken the exercise programme, you will help to improve our understanding of the effectiveness of these different exercise interventions in people with MS. You may find it personally beneficial because you will be able to participate in a 12 week programme of face to face physiotherapy sessions which may improve your balance and mobility. You should understand however that you may not gain benefits from undertaking these exercises.

What are the potential risks of taking part in this study?
In terms of the assessments, the level of ultrasound used for the scanning will be set below the levels recommended by the British Medical Ultrasound Society. This procedure has an excellent record of safety. The researcher undertaking the measurements is trained in the safe and effective use of ultrasound imaging for this specific application.

In terms of the exercise programme you will be prescribed, this will be specifically designed to meet your individual needs and will be closely monitored by your physiotherapist throughout the course of this study. While it is not anticipated that you will experience fatigue, pain or increased spasms while undertaking the exercise programme, nevertheless it is important that you are aware that it is possible that these may occur. Should this happen, it is important that you let your physiotherapist know so that she can modify the exercise programme accordingly, or if necessary that she can withdraw you from the study.

Who is organising and funding the study?
The project is being funded by the Multiple Sclerosis Trust. It is being organized and conducted by Dr Jenny Freeman, Reader in Physiotherapy within the Faculty of Health at Plymouth University.

Other members of the team include:

- Margaret Gear, Specialist Neurological Physiotherapist, Shetlands NHS Trust
- Dr Alan Hough, Senior Lecturer, Faculty of Health, Plymouth University
- Professor John Zajicek, Consultant Neurologist, Peninsula Medical School, Universities of Plymouth and Exeter
- Esther Fox, Research Fellow, Faculty of Health, Plymouth University

Who has reviewed this study?
This study has been reviewed and approved by the *South West Research Ethics Committee.*

What if something goes wrong?
If you wish to complain, or have any concerns about this study then in the first instance please contact the researcher whose details are at the end of the Information Sheet. The Plymouth Guild of Voluntary Service are also there to help, and are available via phone telephone 01752 211818. The normal National Health Service complaints mechanisms should also be available to you.

In the unlikely event that you may be harmed by taking part in this research there are no special compensation arrangements. If you are harmed due to someone's negligence, then you may have grounds for a legal action.

How will I hear about the results of the study?
We anticipate that it will take approximately 18 months for the study to be completed. At the end of this period, if you wish, we will send you a summary of the results of this study. A summary of the results will also be made available on the MS Trust web-site www.mstrust.org.uk

Your rights
Your participation in this study is entirely voluntary. You may withdraw at any time without it affecting your current or future medical treatment in any way. If you agree to take part in this study, you will need to sign a consent form.

Contact for further information
If you require any further information about this study, or have any questions please contact either Esther Fox on 01752 587599 or Dr Jenny Freeman on 01752 588835 during office hours.

Thank you for reading this Information Sheet and considering taking part in the study. If you decide to participate in this study you will be given a copy of this Information Sheet and a signed consent form to keep.

Appendix 4: Table 1: Results of all analyses performed for Pilates vs Relaxation at 12 weeks

(statistically significant between group differences highlighted in yellow, table continued over page).

Outcome measure	Complete case data		Complete case data with outliers removed		LOCF relapses removed		LOCF outliers and relapses removed	
Pilates vs relaxation 12 weeks	**Pilates**	**Relax**	**Pilates**	**relax**	**Pilates**	**relax**	**Pilates**	**Relax**
10 metre timed walk	**n=28**	**n=21**	**n=28**	**n=21**	**n=33**	**n=29**	**n=31**	**n=27**
mean (change score in seconds)	2.03	0.90	2.03	0.90	1.72	0.69	1.74	1.29
standard deviation of change score	3.49	4.04	3.49	4.04	3.29	3.44	2.95	2.60
mean difference with relaxation (seconds)	1.13		1.13		1.03		0.45	
p value	0.30		0.30		0.23		0.55	
lower 95% CI	-1.04		-1.04		-0.68		-1.03	
upper 95% CI	3.30		3.30		2.75		1.92	
Walking Velocity	**n= 28**	**n=21**	**n=28**	**n=21**	**n=33**	**n=29**	**n=31**	**n= 28**
mean (change score in m/s)	0.11	0.04	0.11	0.04	0.10	0.01	0.07	0.03
standard deviation of change score (m/s)	0.19	0.11	0.02	0.11	0.18	0.13	0.15	0.10
mean difference with relaxation (m/s)	0.07		0.74		0.08		0.04	
p value	0.10		0.11		0.04		0.21	
lower 95% CI	-0.01		-0.02		0.00		-0.03	
upper 95% CI	0.16		0.17		0.16		0.11	
Forward Functional Reach	**n=28**	**n=23**	**n=28**	**n=22**	**n=33**	**n=28**	**n= 32**	**n=26**
mean (change score in cm)	3.66	0.31	3.65	1.60	3.10	minus 0.01	2.66	1.57
standard deviation of change score (cm)	4.59	7.95	4.59	5.04	4.42	7.18	3.70	3.79
mean difference with relaxation (cm)	3.34		2.04		3.11		1.09	
p value	0.07		0.14		0.04		0.27	
lower 95% CI	-0.23		-0.72		0.11		-0.89	
upper 95% CI	6.92		4.79		6.12		3.07	

Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.

Outcome measure	Complete case data		Complete case data with outliers removed		LOCF relapses removed		LOCF outliers and relapses removed	
Pilates vs relaxation 12 weeks	**Pilates**	**Relax**	**Pilates**	**relax**	**Pilates**	**relax**	**Pilates**	**Relax**
Lateral Functional Reach	**n=28**	**n= 23**	**n=26**	**n=22**	**n=32**	**n=27**	**n=30**	**n=24**
mean (change score in cm)	2.45	minus 0.53	2.49	-0.01	2.15	minus 0.84	2.17	0.60
standard deviation of change score (cm)	5.46	6.06	4.61	5.65	5.16	5.47	4.38	3.80
mean difference with relaxation (cm)	2.99		2.51		2.98		1.57	
p value	0.07		0.10		0.04		0.17	
lower 95% CI	-0.26		-0.47		0.21		-0.71	
upper 95% CI	6.23		5.49		5.76		3.84	
12 Item Multiple Sclerosis Walking Scale	**n= 26**	**n=23**	**n=25**	**n=23**	**n=31**	**n=29**	**n=30**	**n=28**
mean (change score in points)	9.52	3.52	7.71	3.52	7.99	2.21	6.43	0.68
standard deviation of change score (points)	17.34	13.47	14.98	13.46	16.22	12.39	13.94	9.39
mean difference with relaxation (points)	6.00		4.19		5.77		5.75	
p value	0.19		0.32		0.13		0.07	
lower 95% CI	-3.01		-4.11		-1.73		-0.55	
upper 95% CI	15.01		12.50		13.27		12.05	
Activities Balance Confidence Scale	**n=27**	**n=24**	**n=27**	**n=24**	**n=32**	**n=29**	**n=32**	**n=29**
mean (change score in points)	0.78	0.06	0.78	0.61	0.66	0.07	0.66	0.07
standard deviation of change score (points)	1.36	1.26	1.36	1.26	1.28	1.13	1.28	1.13
mean difference with relaxation (points)	0.72		0.72		0.59		0.59	
p value	0.06		0.06		0.06		0.06	
lower 95% CI	-0.02		-0.02		-0.03		-0.03	
upper 95% CI	1.46		1.46		1.21		1.21	

Perceived difficulty carrying a drink (VAS)	n=27	n=24	n=26	n=24	n=32	n=29	n=29	n=28
mean (change score in points)	0.89	0.17	1.07	0.17	0.75	0.17	0.55	0.28
standard deviation of change score (points)	2.55	2.04	2.39	2.03	2.36	1.85	1.80	1.78
mean difference with relaxation (points)	0.72		0.91		0.58		0.27	
p value	0.27		0.16		0.29		0.58	
lower 95% CI	-0.59		-0.36		-0.52		-0.69	
upper 95% CI	2.03		2.18		1.67		1.22	
Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.								

Table 2: Results of all analyses performed for Pilates vs Relaxation at 16 weeks, table continued over page.

Outcome measure	Complete case data		Complete case data with outliers removed		LOCF relapses removed		LOCF outliers and relapses' removed	
Pilates vs relaxation 16 weeks	**Pilates**	**Relax**	**Pilates**	**relax**	**Pilates**	**relax**	**Pilates**	**Relax**
10 metre timed walk	**n=26**	**n=23**	**n=25**	**n=21**	**n=33**	**n=29**	**n=31**	**n=26**
mean (change score in seconds)	1.51	0.16	1.90	0.11	1.72	0.58	1.74	0.84
standard deviation of change score	3.68	5.08	3.15	4.02	3.29	3.48	2.95	2.05
mean difference with relaxation (seconds)	1.36		1.79		1.14		0.90	
p value	0.29		0.09		0.19		0.20	
lower 95% CI	-1.17		-0.33		-0.58		-0.48	
upper 95% CI	3.88		3.92		2.86		2.27	
Walking Velocity	**n=26**	**n =21**	**n=26**	**n=21**	**n=33**	**n=28**	**n=30**	**n=24**
mean (change score in m/s)	-0.01	- 0.01	0.00	-0.01	0.09	0.01	0.11	0.04
standard deviation of change score	0.12	0.18	0.12	0.18	0.21	0.19	0.17	0.13
mean difference with relaxation (m/s)	0.00		0.00		0.07		0.06	
p value	0.98		0.98		0.16		0.15	
lower 95% CI	-0.09		-0.86		-0.03		-0.02	
upper 95% CI	0.09		0.87		0.18		0.15	
Forward Functional Reach	**n= 26**	**n=24**	**n=26**	**n=22**	**n=33**	**n=28**	**n=33**	**n=27**
mean (change score in cm)	2.14	2.55	2.14	4.15	1.94	1.87	1.94	n=27
standard deviation of change score	7.08	7.80	7.08	5.90	6.41	7.14	6.41	6.38
mean difference with relaxation (cm)	-0.42		-2.01		0.07		-0.58	
p value	0.84		0.30		0.97		0.73	
lower 95% CI	-4.65		-5.83		-3.40		-3.90	
upper 95% CI	3.82		1.82		3.55		2.75	

Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.

Outcome measure	Complete case data		Complete case data with outliers removed		LOCF relapses removed		LOCF outliers and relapses removed	
Pilates vs relaxation 16 weeks	**Pilates**	**Relax**	**Pilates**	**relax**	**Pilates**	**relax**	**Pilates**	**Relax**
Lateral Functional Reach	**n=25**	**n=24**	**n=25**	**n=23**	**n=31**	**n=27**	**n=31**	**n=26**
mean (change score in cm)	1.43	0.38	1.43	1.13	1.12	0.01	1.12	0.66
standard deviation of change score	6.57	7.11	6.57	6.20	5.92	6.46	5.92	5.61
mean difference with relaxation (cm)	1.06		0.87		1.11		0.46	
p value	0.59		0.30		0.50		0.77	
lower 95% CI	-2.88		-3.42		-2.14		-2.62	
upper 95% CI	4.99		4.20		4.37		3.54	
12 Item Multiple Sclerosis Walking Scale	**n=26**	**n=24**	**n=25**	**n=23**	**n=33**	**n=29**	**n=31**	**n=28**
mean (change score in points)	3.57	- 0.10	1.33	-2.17	3.68	- 0.49	3.07	-2.21
standard deviation of change score	21.38	15.55	18.45	12.03	19.72	14.26	16.28	11.05
mean difference with relaxation (points)	3.67		3.50		4.17		5.28	
p value	0.49		0.44		0.35		0.16	
lower 95% CI	-7.04		-5.69		-4.68		-2.05	
upper 95% CI	14.38		12.64		13.03		12.62	
Activities Balance Confidence Scale	**n=25**	**n=25**	**n=24**	**n=25**	**n=32**	**n=29**	**n=30**	**n=28**
mean (change score in points)	0.65	- 0.03	0.87	-0.03	0.61	0.01	0.68	0.10
standard deviation of change score	1.76	1.09	1.39	1.09	1.59	0.99	1.18	0.90
mean difference with relaxation (points)	0.68		0.89		0.59		0.58	
p value	0.11		0.15		0.09		0.04	
lower 95% CI	-0.16		0.18		-0.09		0.02	
upper 95% CI	1.51		1.61		1.28		1.14	

Perceived difficulty carrying a drink (VAS)	n=25	n=24	n=24	n=23	n=32	n=28	n=30	n=27
mean (change score in points)	0.32	-0.42	0.83	-0.17	0.22	- 0.21	- 0.13	0.00
standard deviation of change score	2.34	2.28	2.06	1.99	2.10	1.99	1.63	1.66
mean difference with relaxation (points)	0.74		0.26		0.43		-0.13	
p value	0.27		0.67		0.42		0.76	
lower 95% CI	-0.59		-0.93		-0.63		-1.01	
upper 95% CI	2.07		1.45		1.49		0.74	
Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.								

Table 3: Results of all analyses performed for Standard Exercise vs Relaxation at 12 weeks (significant differences highlighted in yellow, table continued over page).

Outcome measure	Complete case data		Complete case data with outliers removed		LOCF relapses removed		LOCF outliers and relapses removed	
Standard Exercise vs relax at week 12	**SE**	**Relax**	**SE**	**relax**	**SE**	**relax**	**SE**	**Relax**
10 metre timed walk	**n=30**	**n=21**	**n=30**	**n=21**	**n=32**	**n=29**	**n= 31**	**n=27**
mean (change score in seconds)	2.26	0.90	2.26	0.90	2.12	0.69	1.93	1.29
standard deviation of change score	2.24	4.04	2.24	4.04	2.23	3.44	1.99	2.60
mean difference with relaxation (seconds)	1.36		1.35		1.43		0.64	
p value	0.13		0.13		0.05		0.30	
lower 95% CI	-0.42		-0.41		-0.04		-0.57	
upper 95% CI	3.13		3.13		2.90		1.85	
Walking Velocity	**n=30**	**n=21**	**n=30**	**n=21**	**n= 32**	**n=29**	**n= 28**	**n= 28**
mean (change score in m/s)	0.18	0.04	0.18	0.04	0.17	0.01	0.13	0.03
standard deviation of change score	0.16	0.11	0.16	0.11	0.16	0.13	0.13	0.10
mean difference with relaxation (m/s)	0.14		0.14		0.16		0.10	
p value	p<0.01		p<0.01		p<0.01		p<0.01	
lower 95% CI	0.07		0.07		0.08		0.04	
upper 95% CI	0.22		0.22		0.230		0.17	
Forward Functional Reach	**n=30**	**n=23**	**n=27**	**n=22**	**n=31**	**n=28**	**n=28**	**n=26**
mean (change score in cm)	4.59	0.31	4.06	1.60	4.44	- 0.01	3.92	1.57
standard deviation of change score	7.04	7.95	4.87	5.04	6.97	7.18	4.85	3.79
mean difference with relaxation (cm)	4.28		2.45		4.45		2.35	
p value	0.04		0.09		0.02		0.05	
lower 95% CI	0.14		-0.04		0.76		-0.04	
upper 95% CI	8.42		5.32		8.15		4.74	

Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.

Outcome measure	Complete case data		Complete case data with outliers removed		LOCF relapses removed		LOCF outliers and relapses removed	
Standard Exercise vs relax at week 12	**SE**	**Relax**	**SE**	**relax**	**SE**	**relax**	**SE**	**Relax**
Lateral Functional Reach	**n=30**	**n= 23**	**n=28**	**n=22**	**n=31**	**n=27**	**n=28**	**n=24**
mean (change score in cm)	3.69	- 0.53	2.86	-0.01	3.57	- 0.84	2.46	0.60
standard deviation of change score	5.27	6.06	4.34	5.65	5.23	5.47	4.03	3.80
mean difference with relaxation (cm)	4.22		2.89		4.40		1.86	
p value	p<0.01		0.05		p<0.01		0.10	
lower 95% CI	1.09		0.04		1.59		-0.34	
upper 95% CI	7.35		5.71		7.22		4.05	
12 Item Multiple Sclerosis Walking Scale	**n=29**	**n=23**	**n=29**	**n=23**	**n=31**	**n=29**	**n=31**	**n=28**
mean (change score in points)	12.48	3.52	12.47	3.52	11.67	2.22	11.67	0.68
standard deviation of change score	12.67	13.47	12.67	13.46	12.63	12.39	12.63	9.39
mean difference with relaxation (points)	8.96		8.95		9.46		10.99	
p value	0.02		0.02		p<0.01		p<0.01	
lower 95% CI	1.65		1.65		2.99		5.14	
upper 95% CI	16.27		16.26		15.93		16.85	
Activities Balance Confidence Scale	**n=29**	**n=24**	**n=28**	**n=24**	**n=31**	**n=29**	**n=31**	**n=29**
mean (change score in points)	1.11	0.06	1.01	0.61	1.03	0.07	0.94	0.07
standard deviation of change score	1.28	1.26	1.19	1.26	1.27	1.13	1.17	1.13
mean difference with relaxation (points)	1.04		0.95		0.96		0.87	
p value	p<0.01		p<0.01		p<0.01		0.01	
lower 95% CI	0.34		0.26		0.34		0.27	
upper 95% CI	1.75		1.63		1.58		1.47	

Perceived difficulty carrying a drink (VAS)	n=30	n=24	n=29	n=24	n=32	n=29	n=30	n=28
mean (change score in points)	0.57	0.17	0.79	0.17	0.53	0.17	0.67	0.28
standard deviation of change score	1.99	2.04	1.58	2.03	1.93	1.85	1.52	1.78
mean difference with relaxation (points)	0.40		0.62		0.36		0.38	
p value	0.47		0.21		0.46		0.38	
lower 95% CI	-0.71		-0.37		-0.61		-0.49	
upper 95% CI	1.51		1.62		1.33		1.25	

Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.

Table 4: Results of all analyses performed for Standard Exercise vs Relaxation at 16 weeks (significant differences highlighted, continued over page).

Outcome measure	Complete case data		Complete case data with outliers removed		LOCF relapses removed		LOCF outliers and relapses removed	
Standard Exercises vs Relax week 16	SE	Relax	SE	relax	SE	relax	SE	Relax
10 metre timed walk	**n=30**	**n=23**	**n=28**	**n=21**	**n=32**	**n=29**	**n=31**	**n=26**
mean (change score in seconds)	- 0.10	0.16	1.56	0.11	2.12	0.58	1.93	0.84
standard deviation of change score	6.87	5.08	2.82	4.02	2.23	3.48	1.99	2.05
mean difference with relaxation (seconds)	-0.26		1.44		1.53		1.09	
p value	0.88		0.15		0.04		0.05	
lower 95% CI	-3.69		-0.52		0.05		0.01	
upper 95% CI	3.17		3.41		3.02		2.16	
Walking Velocity	**n=30**	**n= 21**	**n=29**	**n=21**	**n=32**	**n=28**	**n=29**	**n=24**
mean (change score in m/s)	- 0.09	- 0.01	-0.89	-0.01	0.09	0.01	0.07	0.04
standard deviation of change score	0.19	0.18	0.19	0.18	0.23	0.19	0.17	0.13
mean difference with relaxation (m/s)	-0.08		-0.08	-0.08	0.07		0.03	
p value	0.12		0.14		0.19		0.49	
lower 95% CI	-0.19		-0.18		-0.04		-0.05	
upper 95% CI	0.02		0.03		0.19		0.11	
Forward Functional Reach	**n= 30**	**n=24**	**n=28**	**n=22**	**n=31**	**n=28**	**n=28**	**n=27**
mean (change score in cm)	4.22	2.55	4.21	4.15	4.09	1.87	3.76	2.52
standard deviation of change score	6.89	7.80	5.17	5.90	6.82	7.15	4.95	6.38
mean difference with relaxation (cm)	1.67		0.06		2.22		1.24	
p value	0.41		0.97		0.22		0.42	
lower 95% CI	-2.35		-3.09		-1.42		-1.84	
upper 95% CI	5.68		3.21		5.86		4.32	

Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.

Outcome measure	Complete case data		Complete case data with outliers removed		LOCF relapses removed		LOCF outliers and relapses removed	
Standard Exercises vs Relax week 16	**SE**	**Relax**	**SE**	**relax**	**SE**	**relax**	**SE**	**Relax**
Lateral Functional Reach	**n=30**	**n=24**	**n=30**	**n=23**	**n=31**	**n=27**	**n=31**	**n=26**
mean (change score in cm)	4.86	0.38	4.86	1.13	4.70	0.01	4.70	0.66
standard deviation of change score	5.73	7.11	5.72	6.20	5.70	6.46	5.70	5.61
mean difference with relaxation (cm)	4.48		3.72		4.69		4.04	
p value	0.01		0.03		0.01		0.01	
lower 95% CI	0.97		0.42		1.49		1.02	
upper 95% CI	7.98		7.02		7.89		7.05	
12 Item Multiple Sclerosis Walking Scale	**n=30**	**n=24**	**n=29**	**n=23**	**n=32**	**n=29**	**n=31**	**n=28**
mean (change score in points)	8.49	-0.1	10.10	-2.17	7.96	- 0.49	9.45	-2.21
standard deviation of change score	15.99	15.55	13.59	12.03	15.60	14.26	13.36	11.05
mean difference with relaxation (points)	8.59		12.27		8.45		11.66	
p value	0.05		p>0.01		0.03		p>0.01	
lower 95% CI	-0.09		5.02		0.77		5.22	
upper 95% CI	17.27		19.52		16.14		18.09	
Activities Balance Confidence Scale	**n=29**	**n=25**	**n=28**	**n=25**	**n=31**	**n=29**	**n= 27**	**n=28**
mean (change score in points)	0.80	-0.03	0.64	-0.03	0.74	0.01	0.52	0.10
standard deviation of change score	1.56	1.09	1.33	1.09	1.52	0.99	0.88	0.90
mean difference with relaxation (points)	0.82		0.66		0.73		0.42	
p value	0.03		0.05		0.03		0.08	
lower 95% CI	0.08		0.00		0.06		-0.06	
upper 95% CI	1.57		1.33		1.40		0.91	

Perceived difficulty carrying a drink (VAS)	n=30	n=24			n=32	n=28	n=31	n=27
mean (change score in points)	0.15	- 0.42	-0.87	-0.17	0.14	- 0.21	-0.08	0.00
standard deviation of change score	2.30	2.28	1.93	1.99	2.22	1.99	1.87	1.66
mean difference with relaxation (points)	0.56		0.88		0.35		-0.08	
p value	0.37		0.09		0.52		0.86	
lower 95% CI	-0.69		-1.01		-0.74		-1.02	
upper 95% CI	1.82		1.18		1.45		0.85	
Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.								

Table 5: Results of all analyses performed for Pilates vs Standard Exercise at 12 weeks (significant differences highlighted, table continued over page).

Outcome measure	Complete case data		Complete case data with outliers removed		LOCF relapses removed		LOCF outliers and relapses removed	
Pilates vs SE week 12	**Pilates**	**SE**	**Pilates**	**SE**	**Pilates**	**SE**	**Pilates**	**SE**
10 metre timed walk	**n=28**	**n=30**	**n=28**	**n=30**	**n=33**	**n=32**	**n=31**	**n= 31**
mean (change score in seconds)	2.03	2.26	2.02	2.26	1.72	2.12	1.74	1.93
standard deviation of change score	3.49	2.24	3.49	2.24	3.29	2.23	2.95	1.99
mean difference between Pilates and SE (seconds)	-0.23		-0.02		-0.40		-0.19	
p value	0.77		0.76		0.57		0.77	
lower 95% CI	-1.79		-1.79		-1.80		-1.47	
upper 95% CI	1.33		1.33		1.00		1.09	
Walking Velocity	**n= 28**	**n=30**	**n=28**	**n=30**	**n=33**	**n= 32**	**n=31**	**n= 28**
mean (change score in m/s)	0.11	0.18	0.11	0.18	0.10	0.17	0.07	0.13
standard deviation of change score	0.19	0.16	0.02	0.16	0.18	0.16	0.15	0.13
mean difference between Pilates and SE (m/s)	-0.07		-0.07		-0.08		-0.06	
p value	0.13		0.13		0.08		0.11	
lower 95% CI	-0.16		-0.16		-0.16		-0.14	
upper 95% CI	0.02		0.02		0.01		0.01	
Forward Functional Reach	**n=28**	**n=30**	**n=28**	**n=27**	**n=33**	**n=31**	**n= 32**	**n=28**
mean (change score in cm)	3.66	4.59	3.65	4.06	3.10	4.44	2.66	3.92
standard deviation of change score	4.59	7.04	4.59	4.87	4.42	6.97	3.70	4.85
mean difference between Pilates and SE (cm)	-0.94		-0.42		-1.34		-1.26	
p value	0.55		0.75		0.36		0.27	
lower 95% CI	-4.09		-2.98		-4.24		-3.52	
upper 95% CI	2.21		2.14		1.55		1.00	

Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.

Outcome measure	Complete case data		Complete case data with outliers removed		LOCF relapses removed		LOCF outliers and relapses removed	
Pilates vs SE week 12	**Pilates**	**SE**	**Pilates**	**SE**	**Pilates**	**SE**	**Pilates**	**SE**
10 metre timed walk	**n=28**	**n=30**	**n=28**	**n=30**	**n=33**	**n=32**	**n=31**	**n= 31**
Lateral Functional Reach	**n=28**	**n=30**	**n=26**	**n=28**	**n=32**	**n=31**	**n=30**	**n=28**
mean (change score in cm)	2.45	3.69	2.49	2.86	2.15	3.57	2.17	2.46
standard deviation of change score	5.46	5.27	4.61	4.34	5.16	5.23	4.38	4.03
mean difference between Pilates and SE (seconds)	-1.23		0.28		-1.42		-0.29	
p value	0.39		0.80		0.28		0.79	
lower 95% CI	-4.06		-2.81		-4.04		-2.51	
upper 95% CI	1.59		2.08		1.20		1.93	
12 Item Multiple Sclerosis Walking Scale	**n= 26**	**n=29**	**n=25**	**n=29**	**n=31**	**n=31**	**n=30**	**n=31**
mean (change score in points)	9.52	12.48	7.71	12.47	7.99	11.67	6.43	11.67
standard deviation of change score	17.34	12.67	14.98	12.67	16.22	12.63	13.94	12.63
mean difference between Pilates and SE (points)	-2.96		-4.76		-3.69		-5.25	
p value	0.47		0.21		0.32		0.13	
lower 95% CI	-11.11		-12.31		-11.07		-12.06	
upper 95% CI	5.20		2.78		3.70		1.57	
Activities Balance Confidence Scale	**n=27**	**n=29**	**n=27**	**n=28**	**n=32**	**n=31**	**n=32**	**n=31**
mean (change score in points)	0.78	1.11	0.78	1.01	0.66	1.03	0.66	0.94
standard deviation of change score	1.36	1.28	1.36	1.19	1.28	1.27	1.28	1.17
mean difference between Pilates and SE (points)	-0.32		-0.22		-0.37		-0.28	
p value	0.37		0.51		0.25		0.38	
lower 95% CI	-1.03		-0.91		-1.02		-0.90	
upper 95% CI	0.39		1.37		0.27		0.35	

Perceived difficulty carrying a drink (VAS)	n=27	n=30	n=26	n=29	n=32	n=32	n=29	n=30
mean (change score in points)	0.89	0.57	1.07	0.79	0.75	0.53	0.55	0.67
standard deviation of change score	2.55	1.99	2.39	1.58	2.36	1.93	1.80	1.52
mean difference between Pilates and SE (points)	0.32		0.28		0.22		-0.11	
p value	0.60		0.60		0.69		0.79	
lower 95% CI	-0.89		-0.81		-0.86		-0.98	
upper 95% CI	1.53		1.37		1.30		0.75	
Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.								

Table 6: Results of all analyses performed for Pilates vs Standard Exercise at 16 weeks (significant differences highlighted continued over page).

Outcome measure	**Complete case data**		**Complete case data with outliers removed**		**LOCF relapses removed**		**LOCF outliers and relapses removed**	
Pilates vs SE week 16	**Pilates**	**SE**	**Pilates**	**SE**	**Pilates**	**SE**	**Pilates**	**SE**
10 metre timed walk	**n=26**	**n=30**	**n=25**	**n=28**	**n=33**	**n=32**	**n=31**	**n=31**
mean (change score in seconds)	1.51	-0.10	1.90	1.56	1.72	2.12	1.74	1.93
standard deviation of change score	3.68	6.87	3.15	2.82	3.29	2.23	2.95	1.99
mean difference between Pilates and SE (seconds)	1.61		0.35		-0.40		-0.19	
p value	0.30		0.67		0.57		0.77	
lower 95% CI	-1.40		-1.30		-1.80		-1.47	
upper 95% CI	4.64		2.00		1.00		1.09	
Walking Velocity	**n=26**	**n=30**	**n=26**	**n=29**	**n=33**	**n=32**	**n=30**	**n=29**
mean (change score in m/s)	-0.01	- 0.09	0.00	-0.89	0.09	0.09	0.11	0.07
standard deviation of change score	0.12	0.19	0.12	0.19	0.21	0.23	0.17	0.17
mean difference between Pilates and SE (m/s)	0.08		0.08		0.00		0.03	
p value	0.06		0.07		0.97		0.46	
lower 95% CI	0.00		-0.01		-0.11		-0.06	
upper 95% CI	0.17		0.17		0.11		0.12	
Forward Functional Reach	**n= 26**	**n= 30**	**n=26**	**n=28**	**n=33**	**n=31**	**n=33**	**n=28**
mean (change score in cm)	2.14	4.22	2.14	4.21	1.94	4.09	1.94	3.76
standard deviation of change score	7.08	6.89	7.08	5.17	6.41	6.82	6.41	4.95
mean difference between Pilates and SE (seconds)	-2.08		-2.07		-2.14		-1.81	
p value	0.27		0.22		0.20		0.23	
lower 95% CI	-5.83		-5.43		-5.45		-4.79	
upper 95% CI	1.66		1.29		1.16		1.16	

Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.

Outcome measure	Complete case data		Complete case data with outliers removed		LOCF relapses removed		LOCF outliers and relapses removed	
Pilates vs SE week 16	**Pilates**	**SE**	**Pilates**	**SE**	**Pilates**	**SE**	**Pilates**	**SE**
Lateral Functional Reach	**n=25**	**n=30**	**n=25**	**n=30**	**n=31**	**n=31**	**n=31**	**n=31**
mean (change score in cm)	1.43	4.86	1.43	4.86	1.12	4.70	1.12	4.70
standard deviation of change score	6.57	5.73	6.57	5.72	5.92	5.70	5.92	5.70
mean difference between Pilates and SE (seconds)	-3.42		-3.42		-3.58		-3.58	
p value	0.04		0.04		0.02		0.02	
lower 95% CI	-6.74		-6.74		-6.53		-6.53	
upper 95% CI	-0.10		-0.10		-0.63		-0.63	
12 Item Multiple Sclerosis Walking Scale	**n=26**	**n=30**	**n=25**	**n=30**	**n=33**	**n=32**	**n=31**	**n=31**
mean (change score in points)	3.57	8.49	1.33	10.10	3.68	7.96	3.07	9.45
standard deviation of change score	21.38	15.99	18.45	13.59	19.72	15.60	16.28	13.36
mean difference between Pilates and SE (points)	-4.92		-8.76		-4.28		-6.37	
p value	0.33		0.05		0.34		0.10	
lower 95% CI	-14.96		-17.53		-13.11		-13.94	
upper 95% CI	5.11		0.01		4.55		1.19	
Activities Balance Confidence Scale	**n=25**	**n=29**	**n=24**	**n=28**	**n=32**	**n=31**	**n=30**	**n= 27**
mean (change score in points)	0.65	0.80	0.87	0.64	0.61	0.74	0.68	0.52
standard deviation of change score	1.76	1.56	1.39	1.33	1.59	1.52	1.18	0.88
mean difference between Pilates and SE (points)	-0.14		0.24		-0.14		0.16	
p value	0.75		0.53		0.72		0.58	
lower 95% CI	-1.05		-0.53		-0.92		-0.40	
upper 95% CI	0.76		0.99		0.65		0.72	

Perceived difficulty carrying a drink (VAS)	n=25	n=30	n=24	n=29	n=32	n=32	n=30	n=31
mean (change score in points)	0.32	0.15	0.83	-0.87	0.22	0.14	-0.13	- 0.08
standard deviation of change score	2.34	2.30	2.06	1.93	2.10	2.22	1.63	1.87
mean difference between Pilates and SE (points)	0.17		0.17		0.08		-0.05	
p value	0.78		0.75		0.88		0.91	
lower 95% CI	-1.09		-0.93		-1.00		-0.95	
upper 95% CI	1.43		1.28		1.16		0.85	
Legend: LOCF= Last observation carried forward, CI= Confidence intervals, VAS= visual analogue scale.								

References

Aggarwal, A., Zutshi, K., Munjal, J., Kumar, S., Sharma, V. (2010) “Comparing stabilization training with balance training in recreationally active individuals,” *International Journal*, 17(5), 244–254

Ainscough-Potts, A.-M., Morrissey, M.C., Critchley, D. (2006) “The response of the transverse abdominis and internal oblique muscles to different postures.,” *Manual therapy*, 11(1), 54–60

Akuthota, V., Nadler, S.F. (2004) “Core strengthening,” *Archives of Physical Medicine and Rehabilitation*, 85(March), 86–92

Alcalde-Cabero, E., Almazán-Isla, J., García-Merino, A., de Sá, J., de Pedro-Cuesta, J. (2013) “Incidence of multiple sclerosis among European Economic Area populations, 1985-2009: the framework for monitoring.,” *BMC neurology*, 13, 58

Allison, G.T., Morris, S.L. (2008) “Transversus abdominis and core stability: has the pendulum swung?,” *British journal of sports medicine*, 42(11), 930–1

Amato, M.P., Portaccio, E. (2007) “Clinical outcome measures in multiple sclerosis.,” *Journal of the neurological sciences*, 259(1-2), 118–22

Arab, A.M., Rasouli, O., Amiri, M., Tahan, N. (2013) “Reliability of ultrasound measurement of automatic activity of the abdominal muscle in participants with and without chronic low back pain.,” *Chiropractic & manual therapies*, 21(1), 37

Asano, M., Dawes, D.J., Arafah, A, Moriello, C., Mayo, N.E. (2009) “What does a structured review of the effectiveness of exercise interventions for persons with

multiple sclerosis tell us about the challenges of designing trials?," *Multiple Sclerosis* 15(4), 412–21

Baert, I., Freeman, J., Smedal, T., Dalgas, U., Romberg, A., Kalron, A., Conyers, H., Elorriaga, I., Gebara, B., Gumse, J., Heric, A., Jensen, E., Jones, K., Knuts, K., Maertens de Noordhout, B., Martic, A., Normann, B., Eijnde, B.O., Rasova, K., Santoyo Medina, C., Truyens, V., Wens, I., Feys, P. (2014) "Responsiveness and Clinically Meaningful Improvement, According to Disability Level, of Five Walking Measures After Rehabilitation in Multiple Sclerosis: A European Multicenter Study.," *Neurorehabilitation and neural repair. epub ahead of print* doi:10.1177/1545968314521010. http://www.ncbi.nlm.nih.gov/pubmed/24503204.

Barker, P.J., Guggenheimer, K.T., Grkovic, I., Briggs, C., Jones, D., CDL, T., PW, H. (2006) "Effects of Tensioning the Lumbar Fasciae on Segmental Stiffness During Flexion and Extension," *Spine*, 31(4), 197–405

Barr, K.P., Griggs, M., Cadby, T. (2007) "Lumbar Stabilization," *American Journal of Physical Medicine & Rehabilitation*, 86(1), 72–80

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

Beales, D.J., O'Sullivan, P.B. and Briffa, N.K. (2009) "Motor control patterns during an active straight leg raise in pain free subjects.," *Spine*, 34(9), 861–70

Behrman, A.L., Light, K.E., Flynn, S.M., Thigpen, M.T. (2002) “Is the functional reach test useful for identifying falls risk among individuals with Parkinson’s disease?,” *Archives of Physical Medicine and Rehabilitation*, 83(4), 538–542

Bengtsson, M., Ohlsson, B., Ulander, K. (2007) “Development and psychometric testing of the Visual Analogue Scale for Irritable Bowel Syndrome (VAS-IBS).,” *BMC gastroenterology*, 7, 16

Bernardo, L. (2007) “The effectiveness of Pilates training in healthy adults: An appraisal of the research literature,” *Journal of Bodywork and Movement Therapies*, 11(2), 106–110

Bethoux, F., Bennett, S. (2011) “Evaluating Walking in Patients with Multiple Sclerosis,” *International Journal of MS Care*, 13(1), 4–14

Bird, M.-L., Fell, J. (2013) “Pilates has Positive Long Term Effects on the Age-Related Decline in Balance and Strength in Older Community Dwelling Men and Women.,” *Journal of Aging Physical Activity*, epub ahead. doi: 10.1123/japa.2013-0006. Epub 2013 Aug 6.

Bird, M.-L., Hill, K.D., Fell, J.W. (2012) “A randomized controlled study investigating static and dynamic balance in older adults after training with Pilates.,” *Archives of physical medicine and rehabilitation*, 93(1), 43–9

Bjerkefors, A., Ekblom, M.M., Josefsson, K., Thorstensson, A. (2010) “Deep and superficial abdominal muscle activation during trunk stabilization exercises with and without instruction to hollow.,” *Manual therapy*, 15(5), 502–7

Bland, J.M., Altman, D.G. (2011) “Comparisons within randomised groups can be very misleading,” *British Medical Journal*, 342, d561

Blaszczyk, J.W. Lowe, D.L. and H.P.D. (1994) “Ranges of postural stability and their changes in the elderly,” *Gait & posture*, 2, 11–17

Bolton, J.E., Wilkinson, R.C. (1998) “Responsiveness of pain scales: a comparison of three pain intensity measures in chiropractic patients.,” *J Manipulative Physiol Ther.*, 21(1), 1–7

Borghuis, J. (2008) “The importance of Sensory-Motor Control in Providing Core Stability: Implications for Measurement and Training,” *Sports Medicine*, 38(11), 893–916

Bowen, J., Gibbons, L., Gianas, A., Kraft, G. (2001) “Self-administered Expanded Disability Status Scale with functional system scores correlates well with a physician-administered test,” *Multiple Sclerosis*, 7(March), 201–206

Brauer, S., Burns, Y., Galley, P. (1999) “Lateral reach: a clinical measure of medio-lateral postural stability.,” *Physiotherapy research international : the journal for researchers and clinicians in physical therapy*, 4(2), 81–8

Briken, S., Gold, S.M., Patra, S., Vettorazzi, E., Harbs, D., Tallner, a, Ketels, G., Schulz, K.H., Heesen, C. (2014) “Effects of exercise on fitness and cognition in progressive MS: a randomized, controlled pilot trial.,” *Multiple Sclerosis* 20(3), 382–90

Brokelman, R.B.G., Haverkamp, D., van Loon, C., Hol, A., van Kampen, A., Veth, R. (2012) “The validation of the visual analogue scale for patient satisfaction after total hip arthroplasty.,” *European orthopaedics and traumatology*, 3(2), 101–105

Brønnum-Hansen, H., Koch-Henriksen, N., Stenager, E. (2004) “Trends in survival and cause of death in Danish patients with multiple sclerosis.,” *Brain*, 127, 844–50

Brooks, C. (2012) “On rethinking core stability exercise programs,” *Australasian Musculoskeletal Medicine*, 9–14

Brown, S.. (2002) “Pilates: where are we now,” *Journal of Dance Medicine & Science*, 6, 108–9

Brown, S.H.M., McGill, S.M. (2010) “A comparison of ultrasound and electromyography measures of force and activation to examine the mechanics of abdominal wall contraction.,” *Clinical biomechanics (Bristol, Avon)*, 25(2), 115–23

Brusse, K.J., Zimdars, S., Zalewski, K.R., Steffen, T.M. (2005) “Research Report Testing Functional Performance in People With Parkinson Disease,” *Physical Therapy*, 85, 134–141.

Burschka, J.M., Keune, P.M., Menge, U., Oy, U.H., Oschmann, P., Hoos, O. (2012) “An exploration of impaired walking dynamics and fatigue in Multiple Sclerosis,” *BMC Neurology*, 12(1), 1

Caldwell, K., Harrison, M., Adams, M., Triplett, N.T. (2009) “Effect of Pilates and taiji quan training on self-efficacy, sleep quality, mood, and physical performance of college students.,” *Journal of bodywork and movement therapies*, 13(2), 155–63

Cameron, M.., Herndon, R.R., Bourdette, D. (2008) “Imbalance in multiple sclerosis: a result of slowed spinal somatosensory conduction,” *Somatosensory & Motor Reserach*, 25(2), 113–122

Cameron, M.H., Wagner, J.M. (2011) “Gait abnormalities in multiple sclerosis: pathogenesis, evaluation, and advances in treatment.,” *Current neurology and neuroscience reports*, 11(5), 507–15

Cattaneo, D., Jonsdottir, J., Repetti, S. (2007) “Reliability of four scales on balance disorders in persons with multiple sclerosis.,” *Disability and rehabilitation*, 29(24), 1920–5

Cattaneo, D., Jonsdottir, J., Zocchi, M., Regola, A. (2007) “Effects of balance exercises on people with multiple sclerosis: a pilot study.,” *Clinical rehabilitation*, 21(9), 771–81.

Cattaneo, D., Regola, A., Meotti, M. (2006) “Validity of six balance disorders scales in persons with multiple sclerosis.,” *Disability and rehabilitation*, 28(12), 789–95

Cetin, N., Bayramoglu, M., Aytar, A., Surenkok, O., Yemisci, O.U. (2008) “Effects of Lower-Extremity and Trunk Muscle Fatigue on Balance,” *The Open Sports Medicine Journal*, 2(1), 16–22

Cholewicki, J. (1997) “Stabilizing function of trunk flexor-extensor muscles a neutral spine posture,” *Spine*, 22(19), 2207–2212

Cholewicki, J., Juluru, K., McGill, S.M. (1999) “Intra-abdominal pressure mechanism for stabilizing the lumbar spine,” *Journal of Biomechanics*, 32(1), 13–17

Cholewicki, J., Polzhofer, G.K., Radebold, a (2000) “Postural control of trunk during unstable sitting.,” *Journal of biomechanics*, 33(12), 1733–7

Cieza, A., Stucki, G. (2008) “The International Classification of Functioning Disability and Health: its development process and content validity,” *Eur J Phys Rehabil Med*, 44(3), 303–313

Coenen, M., Cieza, A., Freeman, J., Khan, F., Miller, D., Weise, A., Kesselring, J. (2011) “The development of ICF Core Sets for multiple sclerosis: results of the International Consensus Conference,” *Journal of Neurology*, 258(8), 1477–1488.

Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciencies* [online], Hillsdale, NJ: Lawrence Erlbaum Associates, available: http://books.google.com/books?hl=en&lr=&id=Tl0N2lRAO9oC&oi=fnd&pg=PR11&dq=Statistical+Power+Analysis+for+the+Behavioral+Sciences&ots=dp6BXlgWUv&sig=ibgAV-xvwwe0QxON9B5PXk4hW4c [accessed 6 Feb 2014].

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), 467–76

Coleman, C., Sobieraj, D., Marinucci, L. (2012) “Minimally important clinical difference of the Timed25-Foot Walk Test: results from a randomized controlled trial in patients with MS,” *Current Medical Research & Opinion*, 1, 49–56.

Collett, J., Dawes, H., Meaney, A., Sackley, C., Barker, K., Wade, D., Izardi, H., Bateman, J., Duda, J., Buckingham, E. (2011) “Exercise for multiple sclerosis: a

single-blind randomized trial comparing three exercise intensities.," *Multiple sclerosis*, 17(5), 594–603

Compston, A., Coles, A. (2008) "Multiple sclerosis.," *Lancet*, 372(9648), 1502–17

Coote, S., Hogan, N., Franklin, S. (2013) "Falls in people with multiple sclerosis who use a walking aid: prevalence, factors, and effect of strength and balance interventions.," *Archives of physical medicine and rehabilitation*, 94(4), 616–21

Costa, L., Maher, C.G., Latimer, J., Smeets, R.J.E.M. (2009) "Ultrasound Imaging for the Measurement of Abdominal Muscle Activity : A Systematic Review," *Physical Therapy*, 89(8)

Costa, L.O.P., Costa, L.C.M., Cancado, R.L., Oliveira, W. and Ferrerira, P.H. (2006) "Short report : Intra-tester reliability of two clinical tests of transversus abdominis muscle recruitment," *Physiotherapy Research International*, 11(1), 48–50.

Crisco, J and Panjabi, M. (1991) "The intersegmental and multisegmental muscles of the lumbar spine. A biomechanical model comparing lateral stabilizing potential.," *Spine*, 16, 793–799.

Critchley, D.J., Coutts, F.J. (2002) "Abdominal Muscle Function in Chronic Low Back Pain Patients," *Physiotherapy*, 88(6), 322–332

Critchley, D.J., Pierson, Z., Battersby, G. (2011) "Effect of pilates mat exercises and conventional exercise programmes on transversus abdominis and obliquus internus abdominis activity: pilot randomised trial.," *Manual therapy*, 16(2), 183–9

Crommert, M.E., Ekblom, M.M., Thorstensson, a (2011) “Activation of transversus abdominis varies with postural demand in standing.,” *Gait & posture*, 33(3), 473–7

Crommert, M.E., Thorstensson, A. (2009) “Trunk muscle reaction to sudden and unexpected pertubations in the absense of upright postural demand,” *Exp Brain Research*, 196, 385–392

Cross, A., Naismith, R. (2014) “Established and novel disease-modifying treatments in multiple sclerosis.,” *J Intern Med*, 275(4), 350–363.

Cruz-Ferreira, A., Fernandes, J., Kuo, Y.-L., Bernardo, L.M., Fernandes, O., Laranjo, L., Silva, A. (2013) “Does pilates-based exercise improve postural alignment in adult women?,” *Women & health*, 53(6), 597–611

Cruz-Ferreira, A., Fernandes, J., Laranjo, L., Bernardo, L.M., Silva, A. (2011) “A Systematic Review of the Effects of Pilates Method of Exercise in Healthy People.,” *Archives of physical medicine and rehabilitation*, 92(12), 2071–2081

Dalgas, U., Severinsen, K., Overgaard, K. (2012) “Relations between 6 minute walking distance and 10 meter walking speed in patients with multiple sclerosis and stroke.,” *Archives of physical medicine and rehabilitation*, 93(7), 1167–72

Dalgas, U., Stenager, E. (2012) “Exercise and disease progression in multiple sclerosis: can exercise slow down the progression of multiple sclerosis?,” *Therapeutic advances in neurological disorders*, 5(2), 81–95

Dalgas, U., Stenager, E., Ingemann-Hansen, T. (2008) “Multiple sclerosis and physical exercise: recommendations for the application of resistance-, endurance- and combined training.,” *Multiple sclerosis* 14(1), 35–53

Danneels, L. A, Cools, A. M., Vanderstraeten, G.G., Cambier, D.C., Witvrouw, E.E., Bourgois, J., de Cuyper, H.J. (2001) “The effects of three different training modalities on the cross-sectional area of the paravertebral muscles.,” *Scandinavian journal of medicine & science in sports*, 11(6), 335–41

Davidson, B.S., Madigan, M.L., Nussbaum, M. a, Wojcik, L. a (2009) “Effects of localized muscle fatigue on recovery from a postural perturbation without stepping.,” *Gait & posture*, 29(4), 552–7

Day, S., Altman, D. (2000) “Blinding in clinical trials and other studies,” *BMJ*, 321(August), 2000

Dayapoglu, N., Tan, M. (2012) “Evaluation of the effect of progressive relaxation exercises on fatigue and sleep quality in patients with Multiple Sclerosis,” *Journal of Alternative and Complemenatry Medicine*, 18(10), 983–987.

DeLuca, G.C., Ebers, G.C., Esiri, M.M. (2004) “Axonal loss in multiple sclerosis: a pathological survey of the corticospinal and sensory tracts.,” *Brain : a journal of neurology*, 127(Pt 5), 1009–18

DeTroyer, A., Estenne, M., Ninane, V., Van Gansbeke, D., Gorini, M. (1990) “Transversus abdominis muscle function in humans,” *Journal of Applied Physiology*, 68(3), 1010–1016.

Dettmers, C., Sulzmann, M., Ruchay-Plössl, A., Gütler, R., Vieten, M. (2009) “Endurance exercise improves walking distance in MS patients with fatigue.,” *Acta Neurol Scand.*, 120(4):251(4), 251–7.

DeWaard, B., Bentrup, B. (2002) “Relationship of the functional reach and lateral reach tests in elderly females,” *Journal of Geriatrics*, 25(3), 2–9

Dickstein, R., Heffes, Y., Laufer, Y., Ben-Haim, Z. (1999) “Activation of selected trunk muscles during symmetric functional activities in poststroke hemiparetic and hemiplegic patients.,” *Journal of neurology, neurosurgery, and psychiatry*, 66(2), 218–21

Dickstein, R., Shefi, S., Marcovitz, E., Villa, Y. (2004) “Electromyographic activity of voluntarily activated trunk flexor and extensor muscles in post-stroke hemiparetic subjects.,” *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 115(4), 790–6

Donner, Allan. and Eliasziw, M. (1987) “Sample size requirements for reliability studies,” *Statisitics in Medicine*, 6, 441–448.

Dorado, C., Calbet, J. a L., Lopez-Gordillo, A., Alayon, S., Sanchis-Moysi, J. (2012) “Marked effects of Pilates on the abdominal muscles: a longitudinal magnetic resonance imaging study.,” *Medicine and science in sports and exercise*, 44(8), 1589–94

Drake, R., Vogl, W., Mitchell, A. (2005) *Grays Anatomy for Students*, 1st ed, Elsevier. pp 54-244

Duncan, P.W., Weiner, D.K., Chandler, J., Studenski, S. (1990) “Functional reach: a new clinical measure of balance.,” *Journal of gerontology*, 45(6), M192–7

Dupont, A.C., Sauerbrei, E.E., Fenton, P. V, Shragge, P.C., Loeb, G.E., Richmond, F.J. (2001) “Real-time sonography to estimate muscle thickness: comparison with MRI and CT.,” *Journal of clinical ultrasound : JCU*, 29(4), 230–6

Von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., Vandenbroucke, J.P. (2007) “[The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies].,” *Revista española de salud pública*, 335(3), 806–8

Van Emmerik, R.E. a, Remelius, J.G., Johnson, M.B., Chung, L.H., Kent-Braun, J. a (2010) “Postural control in women with multiple sclerosis: effects of task, vision and symptomatic fatigue.,” *Gait & posture*, 32(4), 608–14

Endleman, I., Critchley, D.J. (2008) “Transversus abdominis and obliquus internus activity during pilates exercises: measurement with ultrasound scanning.,” *Archives of physical medicine and rehabilitation*, 89(11), 2205–12

English, C., Fisher, L., Thoirs, K. (2012) “Reliability of real-time ultrasound for measuring skeletal muscle size in human limbs in vivo: a systematic review.,” *Clinical rehabilitation*, 26(10), 934–44

English C.K., Thoirs K.A., Fisher L., McLennan, H. and Bernhardt, J. (2012) “Ultrasound is a reliable measure of muscle thickness in acute stroke patients, for some, but not all anatomical sites: A study of intra-rater reliability,” *Ultrasound in Medicie and Biology*, 386-376.

Feinstein, A., Dalgas, U. (2014) “The benefits of exercise in progressive MS: some cautious optimism.,” *Multiple Sclerosis*, 20(3), 269–70

Feinstein, A., Freeman, J., Lo, A.C. (2015) “Treatment of progressive multiple sclerosis: what works, what does not, and what is needed,” *The Lancet Neurology*, 14(2), 194–207

Ferreira, P.H., Ferreira, M.L., Maher, C.G., Refshauge, K., Herbert, R.D., Hodges, P.W. (2010) “Changes in recruitment of transversus abdominis correlate with disability in people with chronic low back pain.,” *British journal of sports medicine*, 44(16), 1166–72

Ferreira, P.H., Ferreira, M.L., Nascimento, D.P., Pinto, R.Z., Franco, M.R., Hodges, P.W. (2011) “Discriminative and reliability analyses of ultrasound measurement of abdominal muscles recruitment.,” *Manual therapy*, 16(5), 463–9

Ferreira, P.H., Ferreira, M.L. and Hodges, P. (2004) “Changes in recruitment of the abdominal muscles in people with low back pain. Ultrasound measurement of muscle activity.,” *Spine*, 29(22), 2560–2566.

Feys, P., Bibby, B., Romberg, A., Santoyo, C., Gebara, B., de Noordhout, B.M., Knuts, K., Bethoux, F., Skjerbæk, A., Jensen, E., Baert, I., Vaney, C., de Groot, V., Dalgas, U. (2014) “Within-day variability on short and long walking tests in persons with multiple sclerosis.,” *Journal of the neurological sciences*, 338(1-2), 183–7

Feys, P., Gijbels, D., Romberg, a, Santoyo, C., Gebara, B., de Noordhout, B.M., Knuts, K., Béthoux, F., de Groot, V., Vaney, C., Dalgas, U. (2012) “Effect of time of day on walking capacity and self-reported fatigue in persons with multiple sclerosis: a multi-center trial.,” *Multiple Sclerosis* 18(3), 351–7

Field, A. (2009) *Discovering Statistics Using SPSS*, 2nd Ed, Sage Publications.pp 153

Filipović-Grčić, P., Matijaca, M., Bilić, I., Džamonja, G., Lušić, I., Čaljkušić, K., Čapkun, V. (2013) “Correlation analysis of visual analogue scale and measures of walking ability in multiple sclerosis patients.,” *Acta Neurol Belg.*, 113(4), 397–402.

Finch, E., Brooks, D., Straford, P., Mayo, N. (2002) “Why are measurement properties important?,” in *Physical Rehabiliation Outcome Measures. A Guide to Enhanced Clinical Decision Making.*, Lippincott Willimans and Wilkins,pp 26–42.

Forouzanfar, T., Weber, W.E.J., Kemler, M., van Kleef, M. (2003) “What Is a Meaningful Pain Reduction in Patients With Complex Regional Pain Syndrome Type 1?,” *The Clinical Journal of Pain*, 19(5), 281–285

Freedman, M.S., Comi, G., De Stefano, N., Barkhof, F., Polman, C.H., Uitdehaag, B.M.J., Lehr, L., Stubinski, B., Kappos, L. (2013) “Moving toward earlier treatment of multiple sclerosis: Findings from a decade of clinical trials and implications for clinical practice,” *Multiple Sclerosis and Related Disorders*, 3(2), 147–155

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

Freeman, J., Fox, E., Gear, M., Hough, A. (2012) “Pilates based core stability training in ambulant individuals with multiple sclerosis: protocol for a multi-centre randomised controlled trial.,” *BMC neurology*, 12(1), 19

Freeman, J., Gear, M., Pauli, A., Cowan, P., Finnigan, C., Hunter, H., Mobberley, C., Nock, A., Sims, R., Thain, J. (2010) “The effect of core stability training on balance and mobility in ambulant individuals with multiple sclerosis: a multi-centre series of single case studies.,” *Multiple Sclerosis* 16(11), 1377–84

Freeman, J.A. (2008) “Chapter 10.4. The patient with degenerative disease: Multiple Sclerosis.,” *In: Pocketbook of Neurological Physiotherapy.* Eds: Lennon, S and Stokes, M. pp 161–175.

Friscia, L.A., Morgan, M.T., Sparto, P.J., Furman, J.M., Whitney, S.L. (2014) “Responsiveness of self-report measures in individuals with vertigo, dizziness, and unsteadiness.,” *Otol Neurotol.*, 35(5), 884–8.

Frzovic, D., Morris, M.E., Vowels, L. (2000) “Clinical tests of standing balance: performance of persons with multiple sclerosis.,” *Archives of physical medicine and rehabilitation*, 81(2), 215–21

Garber, C.E., Blissmer, B., Deschenes, M.R., Franklin, B. a, Lamonte, M.J., Lee, I.-M., Nieman, D.C., Swain, D.P. (2011) “American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise.,” *Medicine and science in sports and exercise*, 43(7), 1334–59

García-Vaquero, M.P., Moreside, J.M., Brontons-Gil, E., Peco-González, N., Vera-Garcia, F.J. (2012) “Trunk muscle activation during stabilization exercises with single and double leg support.,” *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology*, 22, 398–406

Garrett, M., Coote, S. (2009) “Multiple sclerosis and exercise in people with minimal gait impairment – a review,” *Physical Therapy Reviews*, 14(3), 169–180

Garrett, M., Hogan, N., Larkin, A, Saunders, J., Jakeman, P., Coote, S. (2013) “Exercise in the community for people with minimal gait impairment due to MS: an assessor-blind randomized controlled trial.,” *Multiple Sclerosis* 19(6), 782–9

Gatti, R., Corti, M., Bonzani, K., Boccardi, S. (2008) “Electromyographic activity to keep a lower limb in a raised position in healthy subjects and subjects with multiple sclerosis,” *Multiple Sclerosis*, 14(691-693), 691.

Ger, R. (2009) “The clinical anatomy of the anterolateral abdominal wall musculature.,” *Clinical anatomy (New York, N.Y.)*, 22(3), 392–7

Gibbons, S.G.T., Comerford, M.J. (2001) “Strength Versus Stability Part I Strength Versus Stability Part I ; Concept and Terms,” *Orthopaedic Divison Review*, (April), 21–27.

Gijbels, D., Alders, G., Van Hoof, E., Charlier, C., Roelants, M., Broekmans, T., Eijnde, B.O. ’T, Feys, P. (2010) “Predicting habitual walking performance in multiple sclerosis: relevance of capacity and self-report measures.,” *Multiple Sclerosis*, 16(5), 618–26

Gijbels, D., Dalgas, U., Romberg, A., de Groot, V., Bethoux, F., Vaney, C., Gebara, B., Medina, C.., Maamâgi, H., Rasova, K., de Noordhout, B.M., Knuts, K., Feys, P. (2012a) “Which walking capacity tests to use in multiple sclerosis? A multicentre study providing the basis for a core set,” *Multiple Sclerosis* 18(3), 364–7

Gill, N.W., Mason, B.E., Gerber, J.P. (2012) “Lateral abdominal muscle symmetry in collegiate single-sided rowers.,” *International journal of sports physical therapy*, 7(1), 13–9

Gnat, R., Saulicz, E., Miądowicz, B. (2012) “Reliability of real-time ultrasound measurement of transversus abdominis thickness in healthy trained subjects.,” *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, 21(8), 1508–15

Golzari, Z., Fatemeh, S., Soudi, S., Kordi, M., Hashemiet, S. (2010) “Combined exercise training reduces IFN-gamma and IL-17 levels in the plasma and the supernatant of peripheral blood mononuclear cells in women with multiple sclerosis.,” *Int Immunopharmaco*, 10(11), 1415–9.

Goodman, A.D., Brown, T.R., Krupp, L.B., Schapiro, R.T., Schwid, S.R., Cohen, R., Marinucci, L.N., Blight, A.R. (2009) “Sustained-release oral fampridine in multiple sclerosis: a randomised, double-blind, controlled trial.,” *Lancet*, 373(9665), 732–8

Granacher, U., Gollhofer, A., Hortobagyi, T., Kressig, R., Muehlbauer, T. (2013) “The Importance of Trunk Muscle Strength for Balance, Functional Performance and Fall Prevention in Seniors: A Systematic Review,” *Sports Medicine*, 43, 627–641.

Greenhalgh, J., Long, a F., Brettle, a J., Grant, M.J. (1998) “Reviewing and selecting outcome measures for use in routine practice.,” *Journal of evaluation in clinical practice*, 4(4), 339–50

Greenhalgh, T. (2008) *How to Read a Paper, the Basics of Evidence Based Medicine*, 3rd ed, Blackwell Publishing: Oxford. pp101

Grenier, S.G., McGill, S.M. (2007) “Quantification of lumbar stability by using 2 different abdominal activation strategies.,” *Archives of physical medicine and rehabilitation*, 88(1), 54–62

Guclu-Gunduz, A., Citaker, S., Irkec, C., Nazliel, B., Batur-Caglayan, H.Z. (2013) “The effects of pilates on balance, mobility and strength in patients with multiple sclerosis.,” *Neuro Rehabilitation.*, Jul 12(Epub ahead of print).

Gunn, H.J., Newell, P., Haas, B., Marsden, J.F., Freeman, J. a (2013) “Identification of risk factors for falls in multiple sclerosis: a systematic review and meta-analysis.,” *Physical therapy*, 93(4), 504–13

Guralnik, J., Branch, L., Cummings, S., Curb, J. (1989) “Physical performance measures in aging research,” *J Gerontol.*, 44(5), 141–6.

Guyatt, G., Osaba, D., Wu, A., Wyrwich, K., Norma, G. (2002) “Methods to explain the clinical significance of health status measures.,” *Mayo Clinic Proceedings*, 77, 371–83.

Guyatt, G., Walter, S., Norman, G. (1987) “Measuring change over time: Assessing the usefulness of evaluative instruments,” *Journal of Chronic Diseases*, 40(2), 171–178.

Hageman, P. a, Leibowitz, J.M., Blanke, D. (1995) “Age and gender effects on postural control measures.,” *Archives of physical medicine and rehabilitation*, 76(10), 961–5

Haggmark, T., Thorstensson, A. (1979) “Fibre types in human abdominal muscles,” *Acta Physiol Scand*, 104(4), 319–25

Hale, L., Smith, C., Mulligan, H., Treharne, G. (2012) “‘Tell me what you want, what you really really want....’: asking people with multiple sclerosis about enhancing their participation in physical activity,” *Disabil Rehabil 34(22): 1887-1893.*, Disabil Re(22), 1887–1893

Harland, N.J., Dawkin, M.J., Martin, D. (2014) “Relative utility of a visual analogue scale vs a six-point Likert scale in the measurement of global subject outcome in patients with low back pain receiving physiotherapy.,” *Physiotherapy*, 4–8

Hebert, J.R., Corboy, J.R. (2013) “The association between multiple sclerosis-related fatigue and balance as a function of central sensory integration.,” *Gait & posture*, 38(1), 37–42, available: http://www.ncbi.nlm.nih.gov/pubmed/23200463 [accessed 19 Aug 2013].

Hebert, J.R., Corboy, J.R., Manago, M.M., Schenkman, M. (2011) “Effects of vestibular rehabilitation on multiple sclerosis-related fatigue and upright postural control: a randomized controlled trial.,” *Physical therapy*, 91(8), 1166–83

Heesen, C., Böhm, J., Reich, C., Kasper, J., Goebel, M., Gold, S.M. (2008b) “Patient perception of bodily functions in multiple sclerosis: gait and visual function are the most valuable.,” *Multiple Sclerosis*, 14(7), 988–91

Helbostad, J.L., Sturnieks, D.L., Menant, J., Delbaere, K., Lord, S.R., Pijnappels, M. (2010) “Consequences of lower extremity and trunk muscle fatigue on balance and

functional tasks in older people: a systematic literature review.," *BMC geriatrics*, 10, 56

Herrington, L., Davies, R. (2005) "The influence of Pilates training on the ability to contract the Transversus Abdominis muscle in asymptomatic individuals," *Journal of Bodywork and Movement Therapies*, 9(1), 52–57,

Hides, J. a, Richardson, C. a, Jull, G. a (1998) "Use of real-time ultrasound imaging for feedback in rehabilitation," *Manual Therapy*, 3(3), 125–131

Hides, J., Stanton, W., Freke, M., Wilson, S., McMahon, S., Richardson, C. (2008) "MRI study of the size, symmetry and function of the trunk muscles among elite cricketers with and without low back pain.," *British journal of sports medicine*, 42(10), 809–13

Hides, J. Wilson, S and Stanton, W. (2006) "An MRI investigation into the function of the transversus abdominus muscle during 'drawing in' of the abdominal wall.," *Spine*, 31, 175

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), 1509–17

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., Riazi, A., Lamping, D., Fitzpatrick, R., Thompson, A. (2003) “Measuring the impact of MS on walking ability: The 12-item MS Walking Scale (MSWS-12),” *Neurology*, 60, 31–36.

Hobart, J.C., Cano, S.J., Warner, T.T., Thompson, A.J. (2012) “What sample sizes for reliability and validity studies in neurology?,” *Journal of neurology*, 259(12), 2681–9.

Hobart, J.C., D, L., A.J, T. (1996) “NEUROLOGY & PSYCHIATRY Editorial Evaluating neurological outcome measures : the bare essentials,” *Journal of Neurology, Neurosurgery & Psychiatry*, 60, 127–130

Hodges, P., Pengel, L., Herbet, R., Gandevia, S. (2003) “Measurement of muscle contraction with ultrasound imaging.,” *Muscle and Nerve*, 27, 682

Hodges, P., Richardson, C. (1996) “Inefficient muscular stabilisation of the lumbar spine associated with low back pain, a motor control evaluation of transversus abdominis,” *Spine*, 21(22), 2640–50

Hodges, P., Richardson, C. (1999a) “Altered trunk muscle recruitment in people with low back pain with upper limb movement at different speeds.,” *Archives of physical medicine and rehabilitation*, 80(9), 1005–12

Hodges, P., Richardson, C. (1999b) “Transversus abdominis and the superficial abdominal muscles are controlled independently in a postural task,” *Neuroscience letters*, 265, 91–94

Hodges, P.W. (1999) “Is there a role for transversus abdominis in lumbo-pelvic stability?,” *Manual therapy*, 4(2), 74–86

Hoffman, J., Gabel, P. (2013) “Expanding Panjabi’s stability model to express movement: a theoretical model.,” *Medical hypotheses*, 80(6), 692–7

Hoffmann, T.C., Glasziou, P.P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D.G., Barbour, V., Macdonald, H., Johnston, M., Lamb, S.E., Dixon-Woods, M., McCulloch, P., Wyatt, J.C., Chan, a.-W., Michie, S. (2014) “Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide,” *Bmj*, 348(mar07 3), g1687–g1687

Holland, a, O’Connor, R.J., Thompson, a J., Playford, E.D., Hobart, J.C. (2006) “Talking the talk on walking the walk: a 12-item generic walking scale suitable for neurological conditions?,” *Journal of neurology*, 253(12), 1594–602

Hosseini, S., Asl, A., Rostamkhany, H. (2012) “The effect of strength and core stabilization training on physical fitness factors among elderly people,” *World Applied Sciences Journal*, 16(4), 479–484

Hough, A. Marsden, J. Shaw, S. Bakheit, M. Freeman, J. (2009) “Exploring the validity of ultrasound as a measure of trunk activity in acute stroke.,” in *World Congress of Neurology, Bangkok*.

Hu, H., Meijer, O.G., van Dieën, J.H., Hodges, P.W., Bruijn, S.M., Strijers, R.L., Nanayakkara, P.W.B., van Royen, B.J., Wu, W.H., Xia, C. (2011) “Is the psoas a hip flexor in the active straight leg raise?,” *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, 20(5), 759–65

Hu, H., Meijer, O.G., Hodges, P.W., Bruijn, S.M., Strijers, R.L., Nanayakkara, P.W.B., van Royen, B.J., Wu, W., Xia, C., van Dieën, J.H. (2012a) “Understanding the Active Straight Leg Raise (ASLR): An electromyographic study in healthy subjects.,” *Manual therapy*, 1–7

Hu, H., Meijer, O.G., Hodges, P.W., Bruijn, S.M., Strijers, R.L., Nanayakkara, P.W.B., van Royen, B.J., Wu, W.H., Xia, C., van Dieën, J.H. (2012b) “Control of the lateral abdominal muscles during walking.,” *Human movement science*, 31(4), 880–96

Hubley-Kozey, C.L., Hanada, E.Y., Gordon, S., Kozey, J., McKeon, M. (2009) “Differences in abdominal muscle activation patterns of younger and older adults performing an asymmetric leg-loading task.,” *PM & R : the journal of injury, function, and rehabilitation*, 1(11), 1004–13

Hyun, J., Hwangbo, K., Lee, C.-W. (2014) “The effects of pilates mat exercise on the balance ability of elderly females.,” *Journal of physical therapy science*, 26(2), 291–3

Ikezoe, T., Mori, N., Nakamura, M. and Ichihashi, N. (2012) “Effects of age and inactivity due to prolonged bed rest on atrophy of trunk muscles,” *European Journal of Applied Physiology*, 112, 43–48

Irez, G.B., Ozdemir, R.A., Evin, R., Irez, S.G., Korkusuz, F. (2011) “Integrating pilates exercise into an exercise program for 65+ year-old women to reduce falls.,” *Journal of sports science & medicine*, 10(1), 105–11

Iscoe, S. (1998) “Control of abdominal muscles.,” *Progress in neurobiology*, 56(4), 433–506

Ishida, H., Hirose, R., Watanabe, S. (2012) “Comparison of changes in the contraction of the lateral abdominal muscles between the abdominal drawing-in maneuver and breathe held at the maximum expiratory level.,” *Manual therapy*, 17(5), 427–31

Ishida, H., Watanabe, S. (2012) “In influence of inward pressure of the transducer on lateral abdominal thickness during ultrasound imaging,” *Journal of Orthopaedic and Sports Physical Therapy*, 42(9), 815–818

Jacobs, J. V, Henry, S.M., Jones, S.L., Hitt, J.R., Bunn, J.Y. (2011) “A history of low back pain associates with altered electromyographic activation patterns in response to perturbations of standing balance.,” *Journal of neurophysiology*, 106(5), 2506–14

Jagannath, V., Mathew, J.L., Asokan, G. V, Fedorowicz, Z. (2011) “Quality assessment of systematic reviews of health care interventions using AMSTAR.,” *Indian pediatrics*, 48(5), 383–5

John, E.K., Beith, I.D. (2007) “Can activity within the external abdominal oblique be measured using real-time ultrasound imaging?,” *Clinical biomechanics* 22(9), 972–9

Johnson, E., Larsen, a, Ozawa, H., Wilson, C., Kennedy, K. (2007) “The effects of Pilates-based exercise on dynamic balance in healthy adults,” *Journal of Bodywork and Movement Therapies*, 11(3), 238–242

Jonsson, E., Henriksson, M., Hirschfeld, H. (2003) “Does the functional reach test reflect stability limits in elderly people?,” *Journal of rehabilitation medicine*, 35(1), 26–30

Juni P., Altman D.G., and E.M. (2001) “Assessing the quality of controlled clinical trials,” *Bristish Medical Journal*, 323(7303), 42–46.

Kaesler, D., Mellifont, R., Kelly, P., Taaffe, D. (2007) “A novel balance exercise program for postural stability in older adults: A pilot study,” *Journal of Bodywork and Movement Therapies*, 11(1), 37–43

Kage, H., Okuda, M., Nakamura, I., Kunitsugu, I., Sugiyama, S., Hobara, T. (2009) “Measuring methods for functional reach test: comparison of 1-arm reach and 2-arm reach.,” *Archives of physical medicine and rehabilitation*, 90(12), 2103–7

Kaji, A., Sasgawa, S., Kubo, T., Kanehisa, H. (2010) “The transiemt effect of core stability exercises on postural sway during quiet standing,” *Journal of Strength & Conditioning Research*, 24(2), 382–388.

Kaneko, H., Sato, H., Maruyama, H. (2006) “Evaluation of lateral abdominal muscle activity during expiratory threshold loading by ultrasonography,” *Journal of Physical Therapy Science*, 187–191

Kaneko, H. Sato, H. and Maruyama, H. (2005) “Reliability of lateral abdominal muscle thickness measurement using ultrasonography. (abstract only),” *Rigakuryoho Kagaku*, 20(3), 97–201

Kang, K.-Y., Choi, J.-H., Lee, S.-B. (2012) “Effect of Core Strengthening Exercise Programs on Symmetric Double Limb Support and Balance Ability for the Elderly,” *Journal of International Academy of Physical Therapy Research*, 3(1), 378–382

Karatas, M., Cetin, N., Bayramoglu, M., Dilek, A. (2004) “Trunk muscle strength in relation to balance and functional disability in unihemispheric stroke patients.,”

American journal of physical medicine & rehabilitation / Association of Academic Physiatrists, 83(2), 81–7

Karpatkin, H. (2005) “ Multiple Sclerosis and Exercise: a review of the evidence,” *International Journal of Multiple Sclerosis Care*, 7, 36–41

Katz-Leurer, M., Fisher, I., Neeb, M., Schwartz, I., Carmeli, E. (2009) “Reliability and validity of the modified functional reach test at the sub-acute stage post-stroke.,” *Disabil Rehabil.*, 31(3), 243–8.

Kavcic, N., Grenier, S., McGill, S.M. (2004) “Determining the stabilizing role of individual torso muscles during rehabilitation exercises.,” *Spine*, 29(11), 1254–65

Kelly, A.M. (2001) “The minimum clinically significant difference in visual analogue scale pain score does not differ with severity of pain,” *Emerg Med J*, 18, 205–207

Kempen, J.C.E., 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), 980–90

Kesselring, J. (2010) “Multiple Sclerosis Disease Progression in Multiple Sclerosis I . Impaired Mobility and Its Impact on Limitations of Activities and Social Participation,” *European Neurological Review*, 5(1), 56–68

Key, J. (2013) “‘The core’: Understanding it, and retraining its dysfunction,” *Journal of Bodywork and Movement Therapies*, e-pub ahea, 1–19

Kibler, W. Ben, Press, J., Sciascia, A. (2006) “The role of core stability in athletic function.,” *Sports medicine*, 36(3), 189–98

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), 639–46

Kieseier, B.C., Pozzilli, C. (2012) “Assessing walking disability in multiple sclerosis.,” *Multiple sclerosis* 18(7), 914–24

Kiesel, K.B., Uhl, T.L., Underwood, F.B., Rodd, D.W., Nitz, A.J. (2007) “Measurement of lumbar multifidus muscle contraction with rehabilitative ultrasound imaging.,” *Manual therapy*, 12(2), 161–6

Kjølhede, T., Vissing, K., Dalgas, U. (2012) “Multiple sclerosis and progressive resistance training: a systematic review.,” *Multiple sclerosis (Houndmills, Basingstoke, England)*, 18(9), 1215–28

Kloubec, J. (2010) “Pilates for Improvement of Muscle Endurance, Flexibility, Balance, and Posture,” *Journal of Strength & Conditioning Research*, 24(3), 661–7.

Kloubec, J. (2011) “Pilates : how does it work and who needs it ?,” *Muscles, Ligaments and Tendons Journal*, 1(2), 61–66.

Koppenhaver, S.L., Hebert, J.J., Fritz, J.M., Parent, E.C., Teyhen, D.S., Magel, J.S. (2009) “Reliability of rehabilitative ultrasound imaging of the transversus abdominis and lumbar multifidus muscles.,” *Archives of physical medicine and rehabilitation*, 90(1), 87–94

Kordi, R., Rostami, M., Noormohammadpour, P., Mansournia, M.A. (2011) “The effect of food consumption on the thickness of abdominal muscles, employing

ultrasound measurements.," *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, 20(8), 1312–7

Kottner, J., Audigé, L., Brorson, S., Donner, A., Gajewski, B.J., Hróbjartsson, A., Roberts, C., Shoukri, M., Streiner, D.L. (2011) "Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed.," *Journal of clinical epidemiology*, 64(1), 96–106

Kragt, J.J., van der Linden, F. a H., Nielsen, J.M., Uitdehaag, B.M.J., 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), 594–8

Krishnan, V., Kanekar, N., Aruin, A.S. (2012) "Feedforward postural control in individuals with multiple sclerosis during load release.," *Gait & posture*,ebub ahead of print, available: http://www.ncbi.nlm.nih.gov/pubmed/22483546 [accessed 24 May 2012].

Kumar, S. (2010) "EMG in rotation-flexion of the torso," *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology*, 20(6), 1146–54

Kurtzke, J.. (1983) "Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS)," *Neurology*, 33(11), 1444–52

Kuypers, H and Brinkman, J. (1970) "Precentral projections to the different parts of the spinal intermediate zone in teh rhesus monkey," *Brain Research*, 24, 29–48

Laerhoven, H., Zaag-Loonen, H., Derkx, B. (2007) “A comparison of Likert scale and visual analogue scales as response options in children’s questionnaires,” *Acta Paediatrica*, 93(6), 830–835

Lajoie, Y., Gallagher, S.. (2004) “Predicting falls within the elderly community: comparison of postural sway, reaction time, the Berg balance scale and the Activities-specific Balance Confidence (ABC) scale for comparing fallers and non-fallers,” *Archives of Gerontology and Geriatrics*, 38(1), 11–26,

Langdon, D.W., Thompson, A.J. (1999) “Multiple sclerosis : a preliminary study of selected variables affecting rehabilitation outcome,” (June 1998), 94–100.

Lanzetta, D., Cattaneo, D., Pellegatta, D., Cardini, R. (2004) “Trunk control in unstable sitting posture during functional activities in healthy subjects and patients with multiple sclerosis,” *Archives of Physical Medicine and Rehabilitation*, 85(2), 279–283

Lassmann, H. (2011) “Pathophysiology of inflammation and tissue injury in multiple sclerosis: what are the targets for therapy.,” *Journal of the neurological sciences*, 306(1-2), 167–9

Lassmann, H., van Horssen, J., Mahad, D. (2012) “Progressive multiple sclerosis: pathology and pathogenesis.,” *Nature reviews. Neurology*, 8(11), 647–656

Latimer-Cheung, A.E., Pilutti, L. A, Hicks, A.L., Martin Ginis, K., Fenuta, A., Mackibbon, K.A., Motl, R.W. (2013) “The effects of exercise training on fitness, mobility, fatigue, and health related quality of life among adults with multiple sclerosis: a systematic review to inform guideline development.,” *Archives of*

physical medicine and rehabilitation epub ahead of print, available: http://www.ncbi.nlm.nih.gov/pubmed/23669008 [accessed 19 Aug 2013].

Learmonth, Y.C., Paul, L., Miller, L., Mattison, P., McFadyen, A. K. (2011) "The effects of a 12-week leisure centre-based, group exercise intervention for people moderately affected with multiple sclerosis: a randomized controlled pilot study.," *Clinical rehabilitation*,26(7):579-93

Learmonth, Y.C., Dlugonski, D.D., Pilutti, L. a, Sandroff, B.M., Motl, R.W. (2013) "The reliability, precision and clinically meaningful change of walking assessments in multiple sclerosis.," *Multiple Sclerosis* 19(13), 1784–91

Lederman, E. (2010) "The myth of core stability.," *Journal of bodywork and movement therapies*, 14(1), 84–98.

Lee, J.S. (2003) "Clinically Important Change in the Visual Analog Scale after Adequate Pain Control," *Academic Emergency Medicine*, 10(10), 1128–1130

Lesage, F.-X., Berjot, S., Deschamps, F. (2012) "Clinical stress assessment using a visual analogue scale.," *Occupational medicine*, 62(8), 600–5

Liao, C.-F., Lin, S.-I. (2008) "Effects of different movement strategies on forward reach distance.," *Gait & posture*, 28(1), 16–23

Liebenson, C., Karpowicz, A.M., Brown, S.H.M., Howarth, S.J., McGill, S.M. (2009) "The active straight leg raise test and lumbar spine stability.," *PM & R : the journal of injury, function, and rehabilitation*, 1(6), 530–5.

Lim, E.C.W., Poh, R.L.C., Low, A.Y., Wong, W.P. (2011) “Effects of Pilates-based exercises on pain and disability in individuals with persistent nonspecific low back pain: a systematic review with meta-analysis.,” *The Journal of orthopaedic and sports physical therapy*, 41(2), 70–80

Lohr, K.. (2002) “Assessing health status and quality-of-life instruments: Attributes and review criteria,” *Quality of Life Research*, 11, 193–205.

Lublin, F.D., Reingold, S.C. (1996) “Defining the clinical course of Multiple Sclerosis: Results of an international survey,” *Neurology*, 46(4), 907–911.

Lublin, F.D., Reingold, S.C., Cohen, J. a, Cutter, G.R., Sørensen, P.S., Thompson, A.J., Wolinsky, J.S., Balcer, L.J., Banwell, B., Barkhof, F., Bebo, B., Calabresi, P. a, Clanet, M., Comi, G., Fox, R.J., Freedman, M.S., Goodman, A.D., Inglese, M., Kappos, L., Kieseier, B.C., Lincoln, J. a, Lubetzki, C., Miller, A.E.,

Montalban, X., O’Connor, P.W., Petkau, J., Pozzilli, C., Rudick, R. a, Sormani, M.P., Stüve, O., Waubant, E., Polman, C.H. (2014) “Defining the clinical course of multiple sclerosis: the 2013 revisions.,” *Neurology*, 83(3), 278–86

MacDonald, D., Moseley, G., Hodges, P. (2006) “The lumbar multifidus: Does the evidence support clinical beliefs?,” *Mannual Therapy*, 11, 254–263

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, 76–84

Maher, C.G., Sherrington, C., Herbert, R.D., Moseley, A.M., Elkins, M. (2003) “Reliability of the PEDro Scale for Rating Quality of Randomized Controlled Trials,” *Physical Therapy*, 83(8), 713–722

Mannion, A., Pulkovski, N., Schenk, P., Hodges, P., Gerber, H., Loupas, T., Gorelick, M., Sprott, H. (2008) “A new method for the noninvasive determination of abdominal muscle feedforward activity based on tissue velocity information from tissue Doppler imaging.,” *Journal of applied physiology (Bethesda, Md. : 1985)*, 104(4), 1192–201

Mannion, A.F., Pulkovski, N., Toma, V., Sprott, H. (2008) “Abdominal muscle size and symmetry at rest and during abdominal hollowing exercises in healthy control subjects,” *Journal of Anatomy*, 213(2), 173–182

Man-Son-Hing, M., Laupacis, A., O’Rourke, K., Molnar, F.J., Mahon, J., Chan, K.B.Y., Wells, G. (2002) “Determination of the clinical importance of study results.,” *Journal of general internal medicine*, 17(6), 469–76

Mantia, L., Vacchi, L., Rovaris, M., Di Pietrantonj, C., Ebers, G., Fredrikson, S., Filippini, G. (2013) “Interferon β for secondary progressive multiple sclerosis: a systematic review.,” *J Neurol Neurosurg Psychiatry.*, 84(4), 420–6.

Marandi, S., Nejad, V., Shanazari, Z., Zolaktaf, V. (2013) “A comparison of 12 weeks of pilates and aquatic training on the dynamic balance of women with mulitple sclerosis.,” *International Journal Prevenative Medicine*, 4 (Supplement 1), 110–7.

Maranesi, E., Ghetti, G., Rabini, R.A., Fioretti, S. (2014) “Functional reach test: movement strategies in diabetic subjects.,” *Gait & posture*, 39(1), 501–5

Marck, C.H., Hadgkiss, E.J., Weiland, T.J., van der Meer, D.M., Pereira, N.G., Jelinek, G. a (2014) “Physical activity and associated levels of disability and quality of life in people with multiple sclerosis: a large international survey.,” *BMC neurology*, 14(1), 143

Marsden, J.F., Hough, a, Shum, G., Shaw, S., Freeman, J. a (2013) “Deep abdominal muscle activity following supratentorial stroke.,” *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology*, 23(4), 985–90

Martin, C.L., Phillips, B. a, Kilpatrick, T.J., Butzkueven, H., Tubridy, N., McDonald, E., Galea, M.P. (2006) “Gait and balance impairment in early multiple sclerosis in the absence of clinical disability,” *Multiple Sclerosis*, 12(5), 620–628

Mayo, N.E., Bayley, M., Duquette, P., Lapierre, Y., Anderson, R., Bartlett, S. (2013) “The role of exercise in modifying outcomes for people with multiple sclerosis: a randomized trial.,” *BMC neurology*, 13(1), 69

McGalliard, M.K., Dedrick, G.S., Brismée, J.M., Cook, C.E., Apte, G.G., Sizer, P.S. (2010) “Changes in transversus abdominis thickness with use of the abdominal drawing-in maneuver during a functional task.,” *PM & R : the journal of injury, function, and rehabilitation*, 2(3), 187–94

McGill, S. (1991) “Kinetic potential of the lumbar trunk musculature about three orthogonal orthopaedic axes in extreme postures.,” *Spine*, 16(7), 809–15

McGill, S., Cholewicki, J. (2001) “Biomechanical basis for stability: an explanation to enhance clinical utility.,” *J Orthop Sports Phys Ther.*, 31(2), 96–100

McGill, S.M., Karpowicz, A. (2009) “Exercises for spine stabilization: motion/motor patterns, stability progressions, and clinical technique.,” *Archives of physical medicine and rehabilitation*, 90(1), 118–26

McGuigan, C., Hutchinson, M. (2004) “Confirming the validity and responsiveness of the Multiple Sclerosis Walking Scale-12 (MSWS-12).,” *Neurology*, 62(11), 2103–5

McMeeken, J., Beith, I.., Newham, D., Milligan, P., Critchley, D. (2004) “The relationship between EMG and change in thickness of transversus abdominis,” *Clinical Biomechanics*, 19(4), 337

McNeill, W. (2014) “Pilates: ranging beyond neutral - a practical discussion.,” *Journal of bodywork and movement therapies*, 18(1), 124–9

McPherson, S., Watson, T. (2013) “Activation in the Supine Position With Ultrasound Biofeedback Translated to Increased Transversus Abdominis Activation During Upright Loaded Functional Tasks,” *PM&R*, (December), ebup ahead of print available: http://www.ncbi.nlm.nih.gov/pubmed/24316381 [accessed 20 Jan 2014]

Merlo, A., Farina, D., Merletti, R. (2003) “Detection From Surface EMG Signals,” *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, 50(3), 316–323

Meyer-Moock, S., Feng, Y.-S., Maeurer, M., Dippel, F.-W., Kohlmann, T. (2014) “Systematic literature review and validity evaluation of the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC) in patients with multiple sclerosis.,” *BMC neurology*, 14(1), 58

Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P.J., Egger, M., Altman, D.G. (2010) “CONSORT 2010 Explanation and Elaboration:

updated guidelines for reporting parallel group randomised trials," *British Medical Journal*, 340, c869

Mokhtari, M., Nezakatalhossaini, M., Esfarjani, F. (2013) "The Effect of 12-Week Pilates Exercises on Depression and Balance Associated with Falling in the Elderly," *Procedia - Social and Behavioral Sciences*, 70, 1714–1723

Morris, M.E. (2002) "Changes in gait and fatigue from morning to afternoon in people with multiple sclerosis," *Journal of Neurology, Neurosurgery & Psychiatry*, 72(3), 361–365

Morris, S.L., Lay, B., Allison, G.T. (2012) "Corset hypothesis rebutted--transversus abdominis does not co-contract in unison prior to rapid arm movements.," *Clinical biomechanics*, 27(3), 249–54

Morris, S.L., Lay, B., Allison, G.T. (2013) "Transversus abdominis is part of a global not local muscle synergy during arm movement.," *Human movement science*, 32(5), 1176–85

Morton, N.A. De (2009) "The PEDro scale is a valid measure of the methodological quality of clinical trials : a demographic study," *Australian Journal Of Physiotherapy*, 55, 129–133

Motl, R.W., Goldman, M.D., Benedict, R.H.B. (2010) "Walking impairment in patients with multiple sclerosis: exercise training as a treatment option.," *Neuropsychiatric disease and treatment*, 6, 767–74

Motl, R.W., McAuley, E., Snook, E.M. (2005) "Physical activity and multiple sclerosis: a meta-analysis," *Multiple Sclerosis*, 11(4), 459–463

Motl, R.W., 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

Murayama, N., Lin, Y.Y., Salenius, S., Hari, R. (2001a) “Oscillatory interaction between human motor cortex and trunk muscles during isometric contraction.,” *NeuroImage*, 14(5), 1206–13

Murayama, N., Lin, Y.Y., Salenius, S., Hari, R. (2001b) “Oscillatory interaction between human motor cortex and trunk muscles during isometric contraction.,” *NeuroImage*, 14(5), 1206–13

Muscolino, J.E., Cipriani, S. (2004) “Pilates and the ‘powerhouse’—II,” *Journal of Bodywork and Movement Therapies*, 8(2), 122–130

Myers, A., Holliday, P., Harvey, K., Hutchinson, K. (1993) “Functional Performance Measures: Are They Superior to Self-Assessments,” *Journal of Gerontology 48(5): M196-M206.*, 48(5), M196–M206.

Nagar, V., Hooper, T., Dedrick, G., Brismée, J., Sizer, P. (2014) “Effect of recurrent low back pain history on volitional pre-emptive abdominal activation during a loaded functional reach activity.,” *Spine*, 39(2), E89–96.

Newell, D., Shead, V., Sloane, L. (2012) “Changes in gait and balance parameters in elderly subjects attending an 8-week supervised Pilates programme,” *Journal of Bodywork and Movement Therapies*, 6–11

Newsome, S.D., Wang, J.I., Kang, J.Y., Calabresi, P. a, Zackowski, K.M. (2011) “Quantitative measures detect sensory and motor impairments in multiple sclerosis.,” *Journal of the neurological sciences*, 305(1-2), 103–11

Nieves, J., Cosman, F., Herbert, J., Shen, V., Lindsay, R. (1994) “High prevalence of vitamin D deficiency and reduced bone mass in multiple sclerosis.,” *Neurology*, 44, 1687–92

Nilsagård, Y., Carling, A., Forsberg, A. (2012) “Activities-specific balance confidence in people with multiple sclerosis,” *Multiple Sclerosis International*, epub http://dx.doi.org/10.1155/2012/613925

Nilsagard, Y., Lundholm, C., Gunnarsson, L., Denison, E. (2007) “Clinical relevance using timed walk tests and ‘timed up and go’testing in persons with multiple sclerosis,” *Physiotherapy*, 12(2), 105–114

Norasteh, A., Ebrahimi, E., Salavati, M., Rafiei, J., Abbasnejad, E. (2007) “Reliability of B-mode ultrasonography for abdominal muscles in asymptomatic and patients with acute low back pain,” *Journal of Bodywork and Movement Therapies*, 11(1), 17–20

Nuzzo, J.L., Mayer, J.M. (2013) “Body mass normalisation for ultrasound measurements of lumbar multifidus and abdominal muscle size.,” *Manual therapy*, 18(3), 237–42

O’Sullivan, P.B. (2000) “Lumbar segmental ‘instability’: clinical presentation and specific stabilizing exercise management.,” *Manual therapy*, 5(1), 2–12

Örtengren, R., Andersson, G. (1977) “Electromyographic Studies of Trunk Muscles, With Special Reference to the Functional Anatomy of the Lumbar Spine.,” *Spine*, 2(2), 1–88

Owsley, A. (2005) “An introduction to Clinical Pilates,” *Athletic therapy today*, 10(4), 19–25

Le Page, C., Bourdoulous, S., Béraud, E., Couraud, P.O., Rieu, M., A. Ferry (1996) “Effect of physical exercise on adoptive experimental auto-immune encephalomyelitis in rats.,” *Eur J Appl Physiol Occup Physiol.*, 73(1-2), 130–5

Paltamaa, J., Sarasoja, T., Leskinen, 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., Sarasoja, T., Leskinen, E., Wikström, J., Mälkiä, E. (2007) “Measures of physical functioning predict self-reported performance in self-care, mobility, and domestic life in ambulatory persons with multiple sclerosis.,” *Arch Phys Med Rehabil.*, 88(12), 1649–57

Paltamaa, J., Sjögren, T., Peurala, S.H., Heinonen, A. (2012) “Effects of physiotherapy interventions on balance in multiple sclerosis: a systematic review and meta-analysis of randomized controlled trials.,” *Journal of rehabilitation medicine*, 44(10), 811–23

Paltamaa, J., West, H., Sarasoja, T., Wikström, J., Mälkiä, E. (2005) “Reliability of physical functioning measures in ambulatory subjects with MS.,” *Physiotherapy*

research international : the journal for researchers and clinicians in physical therapy, 10(2), 93–109

Panjabi, M.M. (1992) “The stabilising system of the spine part 1: function, dysfunction, adaptation and enhancement.,” *Journal of Spinal Disorders*, 5(4), 383–389

Parkkola, R., Rytökoski, U., Kormano, M. (1993) “Magnetic resonance imaging of the discs and trunk muscles in patients with chronic low back pain and healthy control subjects,” *Spine*, 18(7), 830–6

Pata, R.W., Lord, K., Lamb, J. (2013) “The effect of Pilates based exercise on mobility, postural stability, and balance in order to decrease fall risk in older adults,” *Journal of Bodywork and Movement Therapies*, epub ahead of print available: http://linkinghub.elsevier.com/retrieve/pii/S1360859213001848 [accessed 20 Jan 2014].

Patton, J.L., Pai, Y., Lee, W. a (1999) “Evaluation of a model that determines the stability limits of dynamic balance.,” *Gait & posture*, 9(1), 38–49

Paul, L. Coote, S., Crosbie, J., Dixon, D., Hale, L., Holloway, E., McCrone, P., Miller, L., Saxton, J., Sincock, C., White, L, (2014) “Core outcome measures for exercise studies in people with Multiple Sclerosis: recommendations from a multidisciplinary consensus meeting,” *Multiple Sclerosis*, 20(12), 1641–50.

Pereira, L.M., Marcucci, F.C.I., de Oliveira Menacho, M., Garanhani, M.R., Lavado, E.L., Cardoso, J.R. (2011) “Electromyographic activity of selected trunk muscles in subjects with and without hemiparesis during therapeutic exercise.,” *Journal of*

electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology, 21(2), 327–32

Pereira, L.M., Obara, K., Dias, J.M., Menacho, M.O., Guariglia, D. a, Schiavoni, D., Pereira, H.M., Cardoso, J.R. (2011) “Comparing the Pilates method with no exercise or lumbar stabilization for pain and functionality in patients with chronic low back pain: Systematic review and meta-analysis.,” *Clinical rehabilitation*, Jan;26(1):10-20

Perkin, H.M., Bond, E.A., Thompson, J., Woods, E.C., Smith, C. (2003) “REAL TIME ULTRASOUND : AN OBJECTIVE MEASURE OF SKELETAL MUSCLE,” 99–108.

Petrie, A. and Sabin, C. (2009) *Medical Statistics at a Glance*, 3rd Edition. Wiley-Blackwell.

Petrofsky, J.S., Cuneo, M., Dial, R., Pawley, A.K., Hill, J. (2005) “Core Strengthening and Balance in the Geriatric Population,” *Journal Of Applied Research*, 5(3) 423-33

Petrofsky, J.S., Johnson, E.G., Cuneo, M., Dial, R., Somers, R., Torre, G.D. La, Martinez, A., Forrester, B. (2005) “Abdominal and Lower Back Training for People with Disabilities Using a 6 Second Abs Machine : Effect on Core Muscle Stability,” *Journal Of Applied Research*, 5(2), 345–359

Pike, J., Jones, E., Rajagopalan, K., Piercy, J., Anderson, P. (2012) “Social and economic burden of walking and mobility problems in multiple sclerosis.,” *BMC neurology*, 12(1), 94

Pilutti, L. A, Platta, M.E., Motl, R.W., Latimer-Cheung, A.E. (2014) “The safety of exercise training in multiple sclerosis: A systematic review.,” *Journal of the neurological sciences*, 343(1-2), 3–7

Pilutti, L.A., Dlugonski, D., Sandroff, B.M., Suh, Y., Pula, J.H., Sosnoff, J.J., Motl, R.W. (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

Polman, C.H., Reingold, S.C., Banwell, B., Clanet, M., Cohen, J. a, Filippi, M., Fujihara, K., Havrdova, E., Hutchinson, M., Kappos, L., Lublin, F.D., Montalban, X., O’Connor, P., Sandberg-Wollheim, M., Thompson, A.J., Waubant, E., Weinshenker, B., Wolinsky, J.S. (2011) “Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria.,” *Annals of neurology*, 69(2), 292–302,

Potter, K., Allen, D.D., Bennett, S., Brandfass, K., Cohen, E., Widener, G., Yorke, A. (2012) MS-EDGE Outcome Measures for Research Recommendations for Patients with Multiple Sclerosis [online], *http://www.neuropt.org/docs/degenerative-diseases-sig/ms-edge_research_recs60F4A42650E1.pdf?sfvrsn=2*.

Pretorius, A., Keating, J.L. (2008) “Validity of real time ultrasound for measuring skeletal muscle size,” *Physical Therapy Reviews*, 13(6), 415–426

Preuss, R. A, Grenier, S.G., McGill, S.M. (2005) “Postural control of the lumbar spine in unstable sitting.,” *Archives of physical medicine and rehabilitation*, 86(12), 2309–15

Prosperini, L., Kouleridou, A., Petsas, N., Leonardi, L., Tona, F., Pantano, P., Pozzilli, C. (2011) “The relationship between infratentorial lesions, balance deficit and accidental falls in multiple sclerosis.,” *Journal of the neurological sciences*, 304(1-2), 55–60

Queiroz, B.C., Cagliari, M.F., Amorim, C.F., Sacco, I.C. (2010) “Muscle activation during four Pilates core stability exercises in quadruped position.,” *Archives of physical medicine and rehabilitation*, 91(1), 86–92

Radebold, A., Cholewicki, J., Polzhofer, G., Greene, H. (2001) “Impaired postural control of the lumbar spine is associated with delayed muscle response times in patients with chronic idiopathic low back pain.,” *Spine*, 26(7), 724–30

Raine, S., Meadows, L., Lynch-Ellerington, M. (2009) *Bobath concept:Therory and Clinical Practice in Neurological Rehabilitation*, Wiley Blackwell: UK.

Rankin, G., Stokes, M., Newham, D.J. (2006) “Abdominal muscle size and symmetry in normal subjects.,” *Muscle & nerve*, 34(3), 320–6

Rasouli, O. Arab, A. Amiri, M and Jaberzadeh, S. (2011) “Ultrasound measurement of deep abdominal muscle activity in sitting positions with different stability levels in subjects with and without low back pain.,” *Manual therapy*, 16(4), 338–393

Reeve, A., Dilley, A. (2009) “Effects of posture on thickness of Transversus abdominus in pain-free subjects,” *Manual therapy*, 14(6), 679–684

Richardson, C., Hodges, P., Hides, J. (1999) *Therapeutic Exercise for Lumbopelvic Stabilization*, Churchill Livingstone

Rietberg, M., Brooks, D., Uitdehaag, B., Kwakkel, G. (2004) “Exercise therapy for multiple sclerosis,” *Cochrane Database Syst Rev*, 1(3)

Salhofer-Polanyi, S., Windt, J., Sumper, H., Grill, H., Essmeister, M., Diermayr, G., Zebenholzer, K., Leutmezer, F., Zulehner, G., Vass, K., Asenbaum-Nan, S. (2013)

“Benefits of inpatient multidisciplinary rehabilitation in multiple sclerosis.,” *NeuroRehabilitation*, 33(2), 285–92

Sarasqueta, C., Bergareche, a, Arce, a, Lopez de Munain, a, Poza, J.J., De La Puente, E., Urtasun, M., Emparanza, J.E., Martí Masso, J.F. (2001) “The validity of Hodkinson’s Abbreviated Mental Test for dementia screening in Guipuzcoa, Spain.,” *European journal of neurology : the official journal of the European Federation of Neurological Societies*, 8(5), 435–40

Scheffer, A.C., Schuurmans, M.J., vanDijk, N., van der Hooft, T., de Rooij, S.E. (2010) “Reliability and validity of the visual analogue scale for fear of falling in older persons.,” *Journal of the American Geriatrics Society*, 58(11), 2228–30

Schulz, K.F., Grimes, D.A. (2002) “Epidemiology series Allocation concealment in randomised trials : defending against deciphering,” *The Lancet*, 359, 614–618

Schwid, S.R., Goodman, A.D., McDermott, M.P., C.F.Bever, Cook, S.D. (2002) “Quantitative functional measures in MS: What is a reliable change?,” *Neurology*, 58(8), 1294–1296

Scrimshaw, S. V, Maher, C. (2001) “Responsiveness of visual analogue and McGill pain scale measures.,” *Journal of manipulative and physiological therapeutics*, 24(8), 501–4

Segal, N. A., Hein, J., Basford, J.R. (2004) “The effects of pilates training on flexibility and body composition: An observational study,” *Archives of Physical Medicine and Rehabilitation*, 85(12), 1977–1981

Setts, D. Freund, J. Allison, S and Carpenter, G. (2009) “A rehabilitative ultrasound imaging investigation of lateral abdominal muscle thickness in healthy aging adults,” *Journal of Geriatric Physical Therapy*, 32(2), 16

Shea, B., Hamel, C., Wells, G., Bouter, L., Kristjansson, E., Grimshaw, J., Henry, D., Boers, M. (2009) “AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews.,” *J. Clin Epidemiol.*, 62(10), 1013–20

Shea, B.J., Grimshaw, J.M., Wells, G. A, Boers, M., Andersson, N., Hamel, C., Porter, A.C., Tugwell, P., Moher, D., Bouter, L.M. (2007) “Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews.,” *BMC medical research methodology*, 7, 10

Shea, S., Moriello, G. (2013) “Feasibility and outcomes of a classical Pilates program on lower extremity strength, posture, balance, gait, and quality of life in someone with impairments due to a stroke,” *Journal of Bodywork and Movement Therapies*, epub ahead of print, available: http://linkinghub.elsevier.com/retrieve/pii/S136085921300199X [accessed 20 Jan 2014].

Shumway-Cook, A., Wollacott, M. (2001) “Normal postural control,” in Shumway-Cook A, Wollacott, M., ed., *Motor Control: Theory and Practical Applications.*, Williams and Wilkins, 163–191

Siler, B. (2000) *The Pilates Body*, Micheal Joseph, The Penguin Group.pp 1-13

Silfies, S.P., Squillante, D., Maurer, P., Westcott, S., Karduna, A.R. (2005) “Trunk muscle recruitment patterns in specific chronic low back pain populations.,” *Clinical biomechanics,* 20(5), 465–73

Siqueira Rodrigues, B.G. De, Ali Cader, S., Bento Torres, N.V.O., Oliveira, E.M. De, Martin Dantas, E.H. (2010) “Pilates method in personal autonomy, static balance and quality of life of elderly females.,” *Journal of bodywork and movement therapies*, 14(2), 195–202

Sitilertpisan, P., Pirunsan, U., Puangmali, A., Ratanapinunchai, J., Kiatwattanacharoen, S., Neamin, H., Laskin, J.J. (2011) “Comparison of lateral abdominal muscle thickness between weightlifters and matched controls.,” *Physical therapy in sport : official journal of the Association of Chartered Physiotherapists in Sports Medicine*, 12(4), 171–4

Skjerbæk, a G., Næsby, M., Lützen, K., Møller, a B., Jensen, E., Lamers, I., Stenager, E., Dalgas, U. (2014) “Endurance training is feasible in severely disabled patients with progressive multiple sclerosis.,” *Multiple Sclerosis*, 20(5), 627–30

Smedal, T., Lygren, H., Myhr, K. (2006) “Balance and gait improved in patients with MS after physiotherapy based on the Bobath concept,” *Physiotherapy*, 11(2), 104–116

Smith, C., Olson, K., Hale, L.A., Baxter, D., Schneiders, A.G. (2011) “How does fatigue influence community-based exercise participation in people with multiple sclerosis?,” *Disability and rehabilitation*, 33(23-24), 2362–71

Smith, M.D., Russell, A., Hodges, P.W. (2009) “Do incontinence, breathing difficulties, and gastrointestinal symptoms increase the risk of future back pain?,” *The journal of pain : official journal of the American Pain Society*, 10(8), 876–86

Smith, R.M., Adeney-Steel, M., Fulcher, G., Longley, W. a (2006) “Symptom change with exercise is a temporary phenomenon for people with multiple sclerosis.,” *Archives of physical medicine and rehabilitation*, 87(5), 723–7

Smithson, F., Morris, M.E., Iansek, R. (1998) “Performance on Clinical Tests of Balance in Parkinson ’ s Disease,” *Physical Therapy*, 78, 577–592

Snook, E.M., Motl, R.W. (2009) “Effect of exercise training on walking mobility in multiple sclerosis: a meta-analysis.,” *Neurorehabilitation and neural repair*, 23(2), 108–16

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), 51–5

Sosnoff, J.J., Shin, S., Motl, R.W. (2010) “Multiple sclerosis and postural control: the role of spasticity.,” *Archives of physical medicine and rehabilitation*, 91(1), 93–9

Springer, Barbara. Mielcarek, Billie. Nesfield, T.T. (2006) “Relationship among lateral abdomial muscles, geneder, body mass index and hand dominance,” *Journal of Orthopaedic & Sports Physical Therapy*, 36(5), 289–297

Springer, Barbara., Gill, N. (2007) “Characterization of lateral abdominal muscle thickness in persons with lower extremity amputations,” *Journal Orthopaedic & Sports Physical Therapy*, 10, 635–643

Steffen, T., Seney, M. (2008) “Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism.,” *Phys Ther.*, Jun;88(6), 733–46

Stetts, D., Freund, J., Allison, S. and Carpenter, G. (2009) “A rehabilitative ultrasound imaging investagation of lateral abdominal muscle thickness in healthy aging adults,” *Journal of Geriatric Physical Therapy*, 32(2), 60–66

Stivala, A and Hartley, G. (2014) “The Effects of a Pilates-Based Exercise Rehabilitation Program on Functional Outcome and Fall Risk Reduction in an Aging Adult Status-Post Traumatic Hip Fracture due to a Fall,” *Journal of Geriatric Physical Therapy*, 1, 1–10

Stokes, I. a F., Gardner-Morse, M.G., Henry, S.M. (2011) “Abdominal muscle activation increases lumbar spinal stability: Analysis of contributions of different muscle groups.,” *Clinical biomechanics* 26(8), 797–803

Suri, P., Kiely, D.K., Leveille, S.G., Frontera, W.R., Bean, J.F. (2009) “Trunk muscle attributes are associated with balance and mobility in older adults: a pilot study.,” *PM & R : the journal of injury, function, and rehabilitation*, 1(10), 916–24

Swingler, R.J., Compston, D. a (1992) “The morbidity of multiple sclerosis.,” *The Quarterly journal of medicine*, 83(300), 325–37

Swinnen, E., Beckwée, D., Pinte, D., Meeusen, R., Baeyens, J.-P., Kerckhofs, E. (2012) “Treadmill training in multiple sclerosis: can body weight support or robot

assistance provide added value? A systematic review.," *Multiple sclerosis international*, 2012(2008), 240274

Takayanagi, K.I., Yoshimura, O.S., Nakayama, A.K. (1995) "Sitting Balance and Trunk Muscle Strength in paraplegic patients," Journal of Physical Therapy Science, 7:77-83

Teyhen, D., Gill, N., Whittaker, J., Henry, S., Hides, J., Hodges, P. (2007) "Rehabilitative ultrasound imaging of the abdominal muscles.," *Journal of Orthopaedic and Sports Physical Therapy.*, 27(8), 450–66

Teyhen, D.S., George, S.Z., Dugan, J.L., Williamson, J. (2011) "Inter-Rater Reliability of Ultrasound Imaging of the Trunk Musculature," 347–356

Teyhen, D.S., Rieger, J.L., Westrick, R.B., Miller, A.C., Molloy, J.M., Childs, J.D. (2008) "Changes in deep abdominal muscle thickness during common trunk-strengthening exercises using ultrasound imaging.," *The Journal of orthopaedic and sports physical therapy*, 38(10), 596–605

Teyhen, D.S., Williamson, J.N., Carlson, N.H., Suttles, S.T., O'Laughlin, S.J., Whittaker, J.L., Goffar, S.L., Childs, J.D. (2009) "Ultrasound characteristics of the deep abdominal muscles during the active straight leg raise test.," *Archives of physical medicine and rehabilitation*, 90(5), 761–7

Teyhen D.S., Childs J.D., Stokes M.J., Wright A.C., D.J.L. and G.S.Z. (2012) "Abdominal and Lumbar Multifidus muscle size and symmetry at rest and during contracted states: normative reference ranges," *Journal of ultrasound in medicine*, 31(7), 1099–1110

Thoirs, K., English, C. (2009) “Ultrasound measures of muscle thickness: intra-examiner reliability and influence of body position.,” *Clinical physiology and functional imaging*, 29(6), 440–6

Thomas, J.I., Lane, J. V (2005) “A pilot study to explore the predictive validity of 4 measures of falls risk in frail elderly patients.,” *Archives of physical medicine and rehabilitation*, 86(8), 1636–40

Tinetti, M.E., Richman, D., Powell, L. (1990) “Falls efficacy as a measure of fear of falling.,” *Journal of gerontology*, 45(6), P239–43

Tsao, H., Galea, M.P., Hodges, P.W. (2008) “Concurrent excitation of the opposite motor cortex during transcranial magnetic stimulation to activate the abdominal muscles.,” *Journal of neuroscience methods*, 171(1), 132–9

Tsao, H., Overs, M.E., Wu, J.C.-Y., Galea, M.P., Hodges, P.W. (2008) “Bilateral activation of the abdominal muscles induces longer reaction time.,” *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 119(5), 1147–52

Tulloch, E., Phillips, C., Sole, G., Carman, A., Abbott, J.H. (2012) “DMA Clinical Pilates Directional-Bias Assessment: Reliability and Predictive Validity.,” *The Journal of orthopaedic and sports physical therapy*, 42(8), 676–87

Tyson, S., Connell, L. (2009a) “The psychometric properties and clinical utility of measures of walking and mobility in neurological conditions: a systematic review.,” *Clinical rehabilitation*, 23(11), 1018–33

Tyson, S., Connell, L. (2009b) “How to measure balance in clinical practice. A systematic review of the psychometrics and clinical utility of measures of balance activity for neurological conditions.,” *Clinical rehabilitation*, 23(9), 824–40

Urquhart, D.M., Barker, P.J., Hodges, P.W., Story, I.H., Briggs, C. a (2005) “Regional morphology of the transversus abdominis and obliquus internus and externus abdominis muscles.,” *Clinical biomechanics,* 20(3), 233–41

Urquhart, D.M., Hodges, P.W., Allen, T.J., Story, I.H. (2005) “Abdominal muscle recruitment during a range of voluntary exercises.,” *Manual therapy*, 10(2), 144–53

Urquhart, D.M., Hodges, P.W., Story, I.H. (2005) “Postural activity of the abdominal muscles varies between regions of these muscles and between body positions.,” *Gait & posture*, 22(4), 295–301

Vacha-Haase, T., Thompson, B. (2004) “How to Estimate and Interpret Various Effect Sizes.,” *Journal of Counseling Psychology*, 51(4), 473–481

Van der Linden, M.L., Bulley, C., Geneen, L.J., Hooper, J.E., Cowan, P., Mercer, T.H. (2013) “Pilates for people with multiple sclerosis who use a wheelchair: feasibility, efficacy and participant experiences.,” *Disability and rehabilitation*, 1–8

Vasseljen, O., Fladmark, A.M. (2010) “Abdominal muscle contraction thickness and function after specific and general exercises: a randomized controlled trial in chronic low back pain patients.,” *Manual therapy*, 15(5), 482–9

Vasseljen, O., Fladmark, A.M., Westad, C., Torp, H.G. (2009) “Onset in abdominal muscles recorded simultaneously by ultrasound imaging and intramuscular

electromyography.," *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology*, 19(2), e23–31,

Vasseljen, Ottar, Unsgaard-Tondel, M., Westad, C. and Mork, P.J. (2012) "Effect of core stability exercises on feed-forward activation of deep abdominal muscles in chronic low back pain. A randomized controlled trial," *Spine*, 37(13), 1101–1108.

Vera-Garcia, F. Elvira, J. Brown, S. McGill, S. (2007) "Effects of abdominal stabilization maneuvers on control of spine motion and stability against sudden trunk pertubations," *Journal of Electromyography and Kinesiology*, 17, 556–567

Verheyden, G., Nuyens, G., Nieuwboer, A., Van Asch, P., Ketelaer, P., De Weerdt, W. (2006) "Reliability and validity of trunk assessment for people with multiple sclerosis.," *Physical therapy*, 86(1), 66–76

Wajswelner, H., Metcalf, B., Bennell, K. (2012a) "Clinical pilates versus general exercise for chronic low back pain: randomized trial.," *Medicine and science in sports and exercise*, 44(7), 1197–205

Wajswelner, H., Metcalf, B., Bennell, K. (2012b) "Clinical pilates versus general exercise for chronic low back pain: randomized trial.," *Medicine and science in sports and exercise*, 44(7), 1197–205

Wallden, M. (2013) "The primal nature of core function: in rehabilitation & performance conditioning.," *Journal of bodywork and movement therapies*, 17(2), 239–48

Weir, A., Darby, J., Inklaar, H., Koes, B., Bakker, E., Tol, J.L. (2010) “Core stability: inter- and intraobserver reliability of 6 clinical tests.,” *Clinical journal of sport medicine : official journal of the Canadian Academy of Sport Medicine*, 20(1), 34–8

Wells, C., Kolt, G.S., Bialocerkowski, A. (2012) “Defining Pilates exercise: a systematic review.,” *Complementary therapies in medicine*, 20(4), 253–62

Westad, C., Mork, P.J., Vasseljen, O. (2010) “Location and sequence of muscle onset in deep abdominal muscles measured by different modes of ultrasound imaging.,” *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology*, 20(5), 994–9

Wewers, M.E., Lowe, N.K. (1990) “A Critical Review of Visual Analogue Scales in the Measurement of Clinical Phenomena,” *Reserach in Nursing and Health*, 13, 227–236

White, A.T., Wilson, T.E., Davis, S.L., Petajan, J.H. (2000) “Effect of precooling on physical performance in multiple sclerosis,” *Multiple Sclerosis*, 6, 176–80

White, I.R., Horton, N.J., Carpenter, J., Statistics, R.I.M. a. S., Pocock, S.J. (2011) “Strategy for intention to treat analysis in randomised trials with missing outcome data,” *Bmj*, 342(feb07 1), d40–d40

White, L., Castellano, V. (2008) “Exercise and Brain Health – Implications for Multiple Sclerosis Part II – Immune Factors and Stress Hormones.,” *Sports Med*, 38(3), 179–186

Whitely, E., Ball, J. (2002) “Statistics review 4: Sample size calculations,” *Critical care*, 6(4), 335–341

Willard, F.H., Vleeming, a, Schuenke, M.D., Danneels, L., Schleip, R. (2012) “The thoracolumbar fascia: anatomy, function and clinical considerations.,” *Journal of anatomy*, 221(6), 507–36

Winter, D.A. (1995) “Human balance and posture standing and walking control during.” *Gait and Posture* , 3: 193-214

Withers, G. A Scientific Response to the Article : The Core Stability Myth . [online], *www.ausphysio.com/Files/files-filename-24.pdf*. accessed 27th February 2015, 11:11am

Wood, L., Egger, M., Gluud, L.L., Schulz, K.F., Jüni, P., Altman, D.G., Gluud, C., Martin, R.M., Wood, A.J.G., Sterne, J. a C. (2008) “Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study.,” *BMJ (Clinical research ed.)*, 336(7644), 601–5

Woodward, M. (2005) “Activities-specific and Balance Confidence (ABC) Scale Mini-Mental State Examination (MMSE),” *Clinimetrics*, 51(1998), 1995–1996.

World, H.O. (2001) “The International Classification of Functioning, Disability and Health,” *Geneva*.

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), 144–55

Yang, K.-H., Park, D.-J. (2014) “Reliability of ultrasound in combination with surface electromyogram for evaluating the activity of abdominal muscles in individuals with and without low back pain,” *Journal of Exercise Rehabilitation*, 10(4), 230–235

Yildiz, M. (2012) “The impact of slower walking speed on activities of daily living in patients with multiple sclerosis.,” *International journal of clinical practice*, 66(11), 1088–94

Ytterberg, C., Johansson, S., Andersson, M., Widén Holmqvist, L., von Koch, L. (2008) “Variations in functioning and disability in multiple sclerosis. A two-year prospective study.,” *Journal of neurology*, 255(7), 967–73

Yu, J., Lee, G.C. (2012) “Effect of core stability training using pilates on lower extremity muscle strength and postural stability in healthy subjects,” *Isokinetics and Exercise Science*, 20(2), 141–6

Zaki, R., Bulgiba, A., Ismail, R., Ismail, N.A. (2012) “Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: a systematic review.,” *PloS one*, 7(5), e37908, available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3360667&tool=pmcentrez&rendertype=abstract [accessed 15 Jan 2015].

Zedka, M., Kumar, S., Narayan, Y. (1998) “Electromyographic response of the trunk muscles to postural perturbation in sitting subjects.,” *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology*, 8(1), 3–10

Zisapel, N., Nir, T. (2003) “Determination of the minimal clinically significant difference on a patient visual analog sleep quality scale,” *Journal of sleep research*, 12, 291–298