

Paul, Lorna, Coote, Susan, Crosbie, Jean, Dixon, Diane, Hale, Leigh, Holloway, Ed, McCrone, Paul, Miller, Linda, Saxton, John, Sincock, Caroline, and White, Lesley (2014) Core outcome measures for exercise studies in people with Multiple Sclerosis: recommendations from a multidisciplinary consensus meeting. Multiple Sclerosis. ISSN 1352-4585

Copyright © 2014 The Authors

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

Content must not be changed in any way or reproduced in any format or medium without the formal permission of the copyright holder(s)

When referring to this work, full bibliographic details must be given

<http://eprints.gla.ac.uk/92642>

Deposited on: 24 March 2014

Core outcome measures for exercise studies in people with Multiple Sclerosis:

Recommendations from a multidisciplinary consensus meeting

Lorna Paul¹, Susan Coote², Jean Crosbie³, Diane Dixon⁴, Leigh Hale⁵, Ed Holloway⁶, Paul McCrone⁷, Linda Miller⁸, John Saxton⁹, Caroline Sincock³, Lesley White¹⁰

¹School of Medicine, University of Glasgow, UK, ²Clinical Therapies Department, University of Limerick, Ireland ³Patient representative, ⁴Department of Psychology, University of Strathclyde, UK, ⁵School of Physiotherapy, University of Otago, New Zealand, ⁶Multiple Sclerosis Society, UK, ⁷Institute of Psychiatry, Kings College London, UK, ⁸MS Service NHS Ayrshire and Arran, UK, ⁹School of Rehabilitation Sciences, University of East Anglia, UK, ¹⁰Department of Kinesiology, University of Georgia, USA

Address for correspondence

Dr Lorna Paul

School of Medicine

University of Glasgow

59 Oakfield Avenue

Glasgow G12 8LL

Phone +44 1413306876

Lorna.Paul@glasgow.ac.uk

Key words: Multiple Sclerosis, Outcome Measures, Exercise

Abstract

Background - Evidence shows that exercise is beneficial for people with MS, however statistical pooling of data is difficult due to the diversity of outcome measures used.

Objective- To report the recommendations of an International Consensus Meeting for a core set of outcome measures for use in exercise studies in MS.

Methods – From the 100 categories of the International Classification of Function core sets for MS, 57 categories were considered as likely/potentially likely to be affected by exercise and were clustered into seven core groups. Outcome measures to address each group were evaluated regarding e.g. psychometric properties.

Results - The following are recommended; Modified Fatigue Impact Scale (MFIS) or Fatigue Severity Scale (FSS) for energy and drive, 6 minute walk test (6MWT) for exercise tolerance, timed up and go (TUG) for muscle function and moving around, Multiple Sclerosis Impact Scale (MSIS-29) or MSQoL54 for quality of life and Body Mass Index (BMI) or Waist-hip ratio (WHR) for the health risks associated with excess body fat. A cost effectiveness analysis and qualitative evaluation should be included where possible.

Discussion –Using these core measures ensures that future meta-analyses of exercise studies in MS are more robust and thus more effectively inform practice.

Introduction

While the pharmaceutical industry has made advancements in identifying drug therapies to attenuate symptoms and disease progression for people with MS, great strides have also been made with respect to non-pharmacologic interventions.

Exercise is a safe, non-pharmacological treatment strategy for people with MS that can bring many health benefits, including improvements in muscle power, physical and psychosocial functioning and quality of life¹⁻³. Regular exercise participation may also have an important role to play in the management of fatigue⁴, which negatively affects quality of life⁵, mental alertness⁶ and cognitive processing⁷. Although people with MS have traditionally been advised to refrain from exercise because of the increased risk of triggering an exacerbation of symptoms due a rise in core temperature¹, any increase in symptoms is usually transient (generally 30 minutes or less) and there is no evidence of lasting detrimental effects on fatigue or function⁸.

Indeed, due to a rapidly expanding evidence base, exercise has become one of the key components of rehabilitation and long term management of people with MS to optimise function and maintain health. In the UK, the National Institute for Health and Care Excellence (NICE) recommends aerobic exercise for helping to control the burden of fatigue, motor weakness and musculoskeletal pain in people with MS⁹. This is supported by leading charitable trusts (e.g. the UK MS Society and MS Trust), who are actively promoting the benefits of exercise for improving mobility, mental health status and cardiovascular risk profile to enhance health-related quality of life (HRQoL) and reduce the MS healthcare burden.

Despite the growing evidence base, recent systematic and narrative reviews have highlighted the need for consensus on a core set of outcome measures for exercise studies. A key publication, the 2005 Cochrane Systematic Review¹, reported that statistical pooling

of the data from MS exercise studies was not possible due to the large number of outcome measures used and concluded that there was an urgent need for a consensus on a core set of outcome measures for exercise studies in MS to improve the robustness of future meta-analyses and to ensure that they more effectively inform clinical practice. Since then, a number of reviews on the effects of exercise for people with MS have been published yet the problem of multiple outcome measures and the resultant difficulty of statistical pooling remains.

For example, in their review, Asano *et al.*¹⁰ stated that whilst there was evidence that exercise offered some benefits for people with MS, especially in terms of function and physical activity, advanced statistical analysis was not possible due to the amount and heterogeneity of outcome measures used. Dalgas *et al.*¹¹ reviewed the evidence for resistance, endurance and combined exercise and concluded that there were only a few, methodologically poor and underpowered studies and thus, no conclusions could be made. The review also highlighted the large number of outcome measures included in the studies reviewed. Finally, Motl & Gosney² investigated the effects of exercise on quality of life (QOL) in people with MS and concluded that aerobic exercise in excess of 90 minutes per week can improve QOL but only if measured with a disease specific QOL scale. Importantly, a total of 18 different QOL measures were used in 13 studies.

To advance clinical practice and develop guidelines for the application of exercise for people with MS sufficient evidence of treatment effectiveness is essential. This involves expensive, large scale studies or statistical pooling and meta-analysis of data from small scale studies, which is currently hampered due to the heterogeneity of outcome measures used. Although

the Cochrane Systematic Review in 2005¹ highlighted the need for a core set of outcomes, to date, this has not been proposed. Hence, a meeting of a group of international experts from a range of disciplines was held to discuss outcome measures in MS with the purpose being to recommend a core set of outcome measures for use within exercise studies in MS. This paper reports on the findings of this meeting.

Methods

The consensus group consisted of 12 people, most of whom were internationally recognised for their MS research. They came from a range of professional backgrounds; physiotherapy, exercise science, exercise psychology and health economics as well as a representative from the MS Society UK and two expert patients with MS. Participants attended from the UK, New Zealand, Ireland and the USA. The MS Society UK sent a representative to the meeting and also assisted with recruiting the two individuals with MS through their Research Network. Funding was secured from the Multiple Sclerosis International Federation (MSIF) and the Consensus Meeting took place at the University of Glasgow on 15th and 16th of February 2012.

The EDGE (Evaluation Database to Guide Clinical Effectiveness) Task Force¹² of the American Physical Therapy Association had produced MS-EDGE outcome measures for acute, in - patient and out-patient management of people with MS and also for clinicians undertaking research. While this work primarily focused on clinical practice and specifically on physical therapy, the detailed evaluation by the EDGE group e.g. outlining the validity and reliability of relevant outcome measures was an invaluable resource. Therefore, following dialogue with the Chair of the MS-EDGE group, these documents were used to support/inform some of the discussion of the consensus meeting.

In the preparatory phase discussions took place, by email, in advance of the meeting. The group agreed that an overall framework was required to structure further discussion and for reporting recommendations for a core set of outcome measures. The group agreed that the International Classification of Functioning, Disability and Health (ICF) would be used as the

framework (WHO 2001), specifically the Core Sets for Multiple Sclerosis which had recently been published¹³. The ICF Core Sets for MS contains 40 body functions, 7 body structures and 53 activities and participation categories (100 categories) as well as 38 environmental categories in order to comprehensively describe the function and disability of people with MS based on the ICF. The ICF Core Sets for MS were used as the framework for determining the recommended core set of outcome measures for exercise studies in MS. However, as the environmental categories were part of the external factors, they were not included in the discussion as they could not be influenced by exercise.

At the meeting the group discussed the meaning and value of core outcomes and agreed that the core group of outcome measures should contain as few outcomes as possible but as many as necessary to cover the core domains affected by exercise. In addition the group identified that, for the core set of outcomes to have utility, they would have to be acceptable to various stakeholders: patients and their families, patient groups and charities, clinicians and researchers, research funding agencies, health funders/commissioners, ethics committees and journal editorial boards. The group listed and agreed factors and criteria which would be considered when discussing specific outcome measures (Table 1).

Table 1 Near Here

In Stage 1 of the structured review process, each of the 100 categories from the Core Sets for MS was discussed and classified as a) *likely* b) *potentially likely* or c) *unlikely* to be affected by exercise. As an example, under body functions ‘exercise tolerance’ was classified as likely to be affected by exercise, ‘urinary function’ as potentially affected by exercise and

‘fluency and rhythm of speech functions’ as unlikely to be affected by exercise. Similarly, under Activities and Participation ‘transferring oneself’ was classified as likely to be affected by exercise, ‘higher education’ as potentially affected by exercise and ‘religion and spirituality’ as unlikely to be affected by exercise. The results of Stage 1 are shown in Table 2 which provides the list of categories *likely*, *potentially likely* or *unlikely* to be affected by exercise. Forty categories were considered *likely* to be affected by exercise, 17 as *potentially likely* to be affected by exercise and 36 as *unlikely* to be affected by exercise (Table 2). Those categories *unlikely* to be affected by exercise were then excluded from further discussion. The seven body structures were considered at the next stage.

Table 2 Near Here

In Stage 2, the categories deemed to be *likely* or *potentially likely* to be affected by exercise were then examined and, where relevant, combined into groups. For example 10 categories were clearly related to activities of daily living (ADL) e.g. dressing, eating and were therefore grouped into a category of ADL, 10 categories related to moving around e.g. moving around in different locations, changing body position and four categories were related to muscle function e.g. muscle tone, muscle power.

Stage 2 resulted in eight groupings: energy and drive (fatigue), emotional function, exercise tolerance, muscle function, moving around, ADL, recreation and leisure and quality of life.

Also at this stage consideration was given to the seven body structures. These were mainly anatomical areas e.g. structure of lower limb, brain, trunk and as such likely to be affected

by exercise but were discussed within the groupings identified. These were grouped under body structures, thus making nine groupings.

In Stage 3 each grouping was discussed in terms of whether assessment of that grouping in practice would be determined as a core part of the assessment in studies of exercise in MS. From this discussion, emotional function and recreation and leisure were not deemed to be core and therefore were not included in further discussion.

In Stage 4 the remaining seven groupings were considered in small groups. A list of possible outcome measures was drawn up which would address each grouping. The MS-EDGE documents were used as a resource for this stage. Each outcome measure was considered in relation to the specific criteria previously agreed (Table 1).

In Stage 5 each small group provided a summary of their findings to the whole group. The group discussed and debated the utility of each measure and subsequently agreed on the outcome measure being recommended for each grouping.

In the final stage, Stage 6 the list of core outcome measures was considered by the group and final recommendations agreed.

Results

The consensus group recommend that the outcome measures shown in Table 3 are used, where possible, in studies of exercise in MS. Table 4 shows further detail on the recommended outcomes in terms of the criteria in Table 1.

Table 3 Near Here

Table 4 Near Here

The group recommended that, based on the current published literature available, all exercise based research studies in people with MS should include measures of energy and drive (fatigue), exercise tolerance, muscle function, moving around, ADL, quality of life and body structures. For energy and drive (fatigue) the Modified Fatigue Impact Scale (MFIS)¹⁶ is recommended as it encompasses a multidimensional assessment of fatigue however in situations where unidimensional assessment is sufficient the Fatigue Severity Scale (FSS)¹⁷ could be used. For the assessment of exercise tolerance the 6 minute walk test (6MWT)¹⁸ is recommended. The timed up and go (TUG)¹⁹ is recommended to assess both muscle function and moving around. For quality of life the Multiple Sclerosis Impact Scale (MSIS-29)¹⁴ or MSQoL54¹⁵ are recommended and the simple measures of Body Mass Index (BMI) or Waist-hip ratio (WHR) are recommended as measures of the body composition²⁰.

The group considered a measure of Activities of Daily Living (ADL) to be important however it was not possible to recommend a suitable measure as none of the available measures met the necessary criteria (Table 1).

All the outcome measures recommended are appropriate across a range of disability with the exception of TUG and 6MWT which are not appropriate for those who are non-ambulatory. For those who are non-ambulatory, for the 'moving around' dimension, it was not possible to recommend a measure which was robust, valid and did not require a licence for use.

Although there was unanimous agreement of the importance of including assessments of cost-effectiveness it was not possible to recommend one specific resource use measure due to the differing health care systems. Rather, the principle of taking a comprehensive approach to measuring service use and broader societal costs should be adopted and locally appropriate measures (e.g. patient questionnaires, administrative data) be used for this. In some countries it is appropriate to combine cost data with quality-adjusted life years (QALYs) using a measure such as the EQ5D³⁹. Where QALYs are not required for healthcare decision making this approach may not be required.

The group also recommended that, where appropriate, the results of the quantitative measures recommended be supplemented with a qualitative evaluation. The latter is important to provide an understanding of the obtained results and to glean a more comprehensive perspective of how exercise programmes may affect people with MS.

Discussion

Our goal was to identify and select the most appropriate outcome measures for use in exercise based research as one step in the process toward the application and evaluation of exercise in people with MS.

The outcome measures recommended by the consensus group are a combination of patient reported outcome measures (PROMS); MSIS-29, MSQoL53, MFIS, FSS: rater/time based tests; TUG, 6MWT, and objective measures; BMI, waist/hip ratio. Schaffler et al (2013)⁴⁰ suggest that such a combination is an appropriate triad of outcome measures, and in particular they highlight the growing importance of the inclusion of PROMS in the assessment of MS. Additionally the core outcomes consider a range of body structures, functions, activity limitations and participation restrictions in line with the ICF core sets for MS¹³.

The core outcome measures recommended are commonly used assessments which are relatively quick to complete, require no specialised equipment or training and have minimal cost implications. It should therefore be feasible to include these outcome measures in most studies of exercise in MS. Also of note, with the exception of FSS and the body composition measures (BMI and WHR), all the core outcomes recommended are contained within the MS EDGE Outcome measures for Research (of which there are 27 recommended measures). The omission of the body composition measures (BMI and WHR) from the EDGE measures may be a reflection of this consensus group being multi-professional, including exercise science, compared to the specific physical therapy focus of the EDGE group. The inclusion of body composition measures BMI and WHR, which assess the health risks associated with excess

body fat, is viewed as a strength of the recommended core group of outcome measures, as their inclusion allows the research/practise agenda to move beyond the symptom management/disability focus and to consider the role of exercise in the prevention of secondary health problems such as cardiovascular disease and diabetes in people with MS.

Importantly, the outcome measures recommended are not intended as the only measures for use in exercise studies but that these core measures are included within the battery of relevant outcome measures which address the research questions posed by the investigators of individual studies. For example a study designed to investigate the effects of resistance muscle training would include outcome measures such as muscle strength or power but would also include the recommended core outcome measures.

In our discussion of the two related areas of Activities of Daily Living and 'moving around' for those who were non-ambulatory, we determined that no measure was available which fulfilled our criteria for inclusion as an outcome measure (Table 1). Although there are many ADL measures most require specific training in their use e.g. the Functional Independence Measure (FIM) or require a licence e.g. the Canadian Occupational Performance Measure (COPM). These measures are well validated within the literature however due to accessibility issues are not appropriate within a core set of outcome measures. The difficulty in recommending a specific measure in this area may be a reflection of the literature base being focussed mainly on the effects of exercise in those with relapsing remitting forms of the disease with relatively few studies investigating the effects of exercise for those with progressive MS. There is an urgent need to develop a robust, easily accessible outcome measure to evaluate ADL and 'moving around' in people with MS who are non ambulatory.

We propose that these outcome measures be reviewed regularly, at least every five years. This is important in light of the evolving evidence base, the publication of new outcome measures and the continued validation/update of existing outcome measures. For example the Neurological Fatigue Index for Multiple Sclerosis (NFI-MS)⁴¹ and the Fatigue Scale for Motor and Cognitive Functions (FSMC)⁴² are relatively new measures of fatigue. These were considered but at the time of the consensus meeting there was insufficient evidence on their psychometric properties or their research utility to consider them for inclusion as a core outcome measure⁴¹ However, when the core outcomes are reviewed the results of recent studies^{41, 43, 44} may provide stronger evidence to support their inclusion.

Limitations

The consensus group was relatively small, all English speaking and, although it was multi-disciplinary it did not include representatives from some relevant professional groups such as occupational therapists which would have been valuable. While the group contained internationally leading researchers and expert reviewers for the MS Society budgetary restrictions limited the number of geographical locations represented. It is possible that a different group of researchers might recommend a different set of core measures, however it should be noted that consensus was achieved by those present.

All the core outcome measures are widely used and available in English formats. Future reviews of the core outcome measures should consider if these outcomes are validated in a range of different languages to support the use of the core outcomes in countries where English is not the principal language.

Conclusion

The consensus group recommendations are that all exercise studies in MS include the following outcome measures; the Modified Fatigue Impact Scale (MSIF) or the Fatigue Severity Scale (FSS), the 6 minute walk test (6MWT), the timed up and go (TUG), the Multiple Sclerosis Impact Scale (MSIS-29) or MSQoL54 and Body Mass Index (BMI) or Waist-hip ratio (WHR). These measures should not be to the exclusion of others relevant to the research question(s) of specific studies. Where possible, it is recommended that the core outcome measures are augmented by qualitative evaluation and an assessment of the cost/benefit of the exercise intervention, based on local nationally accepted methods, is included.

Acknowledgements

This work was supported by an International Meeting Grant from the Multiple Sclerosis International Foundation (MSIF). The Multiple Sclerosis Society, UK assisted with recruiting and supporting our patient representatives. We thank Yvonne Learmonth for her invaluable assistance during the meeting.

Conflict of Interest Statement

None Declared

References

1. Rietberg MB, Brooks D, Uitdehaag BM and Kwakkel G. Exercise therapy for multiple sclerosis. *The Cochrane database of systematic reviews*. 2005: CD003980.
2. Motl RW and Gosney JL. Effect of exercise training on quality of life in multiple sclerosis: a meta-analysis. *Multiple sclerosis*. 2008; 14: 129-35.
3. Dalgas U, Stenager E, Jakobsen J, et al. Fatigue, mood and quality of life improve in MS patients after progressive resistance training. *Multiple sclerosis*. 2010; 16: 480-90.
4. Andreassen AK, Stenager E and Dalgas U. The effect of exercise therapy on fatigue in multiple sclerosis. *Multiple sclerosis*. 2011; 17: 1041-54.
5. Amato MP, Ponziani G, Rossi F, Liedl CL, Stefanile C and Rossi L. Quality of life in multiple sclerosis: the impact of depression, fatigue and disability. *Multiple sclerosis*. 2001; 7: 340-4.
6. Weinges-Evers N, Brandt AU, Bock M, et al. Correlation of self-assessed fatigue and alertness in multiple sclerosis. *Multiple sclerosis*. 2010; 16: 1134-40.
7. Andreassen AK, Spliid PE, Andersen H and Jakobsen J. Fatigue and processing speed are related in multiple sclerosis. *European journal of neurology : the official journal of the European Federation of Neurological Societies*. 2010; 17: 212-8.
8. Smith RM, Adeney-Steel M, Fulcher G and Longley WA. Symptom change with exercise is a temporary phenomenon for people with multiple sclerosis. *Archives of physical medicine and rehabilitation*. 2006; 87: 723-7.
9. National Institute for Clinical Excellence. Multiple Sclerosis: management of multiple sclerosis in primary and secondary care. London 2003.
10. Asano M, Dawes DJ, Arafah A, Moriello C and Mayo NE. 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*. 2009; 15: 412-21.
11. Dalgas U, Stenager E and Ingemann-Hansen T. Multiple sclerosis and physical exercise: recommendations for the application of resistance-, endurance- and combined training. *Multiple sclerosis*. 2008; 14: 35-53.
12. Potter K, Cohen ET, Allen DD, et al. Outcome Measures for Individuals With Multiple Sclerosis: Recommendations From the American Physical Therapy Association Neurology Section Task Force. *Physical therapy*. 2013.

13. Coenen M, Cieza A, Freeman J, et al. The development of ICF Core Sets for multiple sclerosis: results of the International Consensus Conference. *Journal of neurology*. 2011; 258: 1477-88.
14. Hobart J, Lamping D, Fitzpatrick R, Riazi A and Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. *Brain : a journal of neurology*. 2001; 124: 962-73.
15. Vickrey BG, Hays RD, Harooni R, Myers LW and Ellison GW. A health-related quality of life measure for multiple sclerosis. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. 1995; 4: 187-206.
16. Fisk JD, Ritvo PG, Ross L, Haase DA, Marrie TJ and Schlech WF. Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 1994; 18 Suppl 1: S79-83.
17. Krupp LB, LaRocca NG, Muir-Nash J and Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of neurology*. 1989; 46: 1121-3.
18. Goldman MD, Marrie RA and Cohen JA. Evaluation of the six-minute walk in multiple sclerosis subjects and healthy controls. *Multiple sclerosis*. 2008; 14: 383-90.
19. Podsiadlo D and Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *Journal of the American Geriatrics Society*. 1991; 39: 142-8.
20. *ACSM Guidelines for Exercise Testing and Prescription*. 9th edition ed.: Lippincott Williams and Watkins, 2013.
21. Riazi A, Hobart JC, Lamping DL, Fitzpatrick R and Thompson AJ. Multiple Sclerosis Impact Scale (MSIS-29): reliability and validity in hospital based samples. *Journal of neurology, neurosurgery, and psychiatry*. 2002; 73: 701-4.
22. Costelloe L, O'Rourke K, Kearney H, et al. The patient knows best: significant change in the physical component of the Multiple Sclerosis Impact Scale (MSIS-29 physical). *Journal of neurology, neurosurgery, and psychiatry*. 2007; 78: 841-4.
23. Ozakbas S, Akdede BB, Kosehasanogullari G, Aksan O and Idiman E. Difference between generic and multiple sclerosis-specific quality of life instruments regarding the assessment of treatment efficacy. *Journal of the neurological sciences*. 2007; 256: 30-4.

24. Amtmann D, Bamer AM, Noonan V, Lang N, Kim J and Cook KF. Comparison of the psychometric properties of two fatigue scales in multiple sclerosis. *Rehabilitation psychology*. 2012; 57: 159-66.
25. Kos D, Kerckhofs E, Nagels G, et al. Assessing fatigue in multiple sclerosis: Dutch modified fatigue impact scale. *Acta neurologica Belgica*. 2003; 103: 185-91.
26. Learmonth YC, Dlugonski DD, Pilutti LA, Sandroff BM and Motl RW. The reliability, precision and clinically meaningful change of walking assessments in multiple sclerosis. *Multiple Sclerosis Journal*. 2013; 19: 1784-91.
27. Rietberg MB, Van Wegen EE and Kwakkel G. Measuring fatigue in patients with multiple sclerosis: reproducibility, responsiveness and concurrent validity of three Dutch self-report questionnaires. *Disability and rehabilitation*. 2010; 32: 1870-6.
28. Toomey E and Coote S. Between-Rater Reliability of the 6-Minute Walk Test, Berg Balance Scale, and Handheld Dynamometry in People with Multiple Sclerosis. *International Journal of MS Care*. 2013; 15: 1-6.
29. Paltamaa J, West H, Sarasoja T, Wikström J and Mälkiä E. Reliability of physical functioning measures in ambulatory subjects with MS. *Physiotherapy Research International*. 2005; 10: 93-109.
30. Learmonth YC, Paul L, McFadyen AK, Mattison P and Miller L. Reliability and clinical significance of mobility and balance assessments in multiple sclerosis. *International journal of rehabilitation research Internationale Zeitschrift fur Rehabilitationsforschung Revue internationale de recherches de readaptation*. 2012; 35: 69-74.
31. Fry DK and Pfalzer LA. Reliability of Four Functional Tests and Rating of Perceived Exertion in Persons with Multiple Sclerosis. *Physiotherapy Canada*. 2006; 58: 212-20.
32. Paltamaa J, Sarasoja T, Leskinen E, Wikstrom J and Malkia E. Measuring deterioration in international classification of functioning domains of people with multiple sclerosis who are ambulatory. *Physical therapy*. 2008; 88: 176-90.
33. Nilsagard Y, Lundholm C, Gunnarsson L and Denison E. Clinical relevance using timed walk tests and 'timed up and go' testing in persons with multiple sclerosis. *Physiotherapy research international*. 2007; 12: 105-14.
34. Kragt JJ, Nielsen JM, van der Linden FA, Polman CH and Uitdehaag BM. Disease progression in multiple sclerosis: combining physicians' and patients' perspectives? *Multiple sclerosis*. 2011; 17: 234-40.
35. Grace J, Mendelsohn A and Friedman JH. A comparison of fatigue measures in Parkinson's disease. *Parkinsonism & related disorders*. 2007; 13: 443-5.

36. Steffen TM, Hacker TA and Mollinger L. Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. *Physical therapy*. 2002; 82: 128-37.
37. Teramoto S, Ohga E, Ishii T, Yamaguchi Y, Yamamoto H and Mastsuse T. Reference value of six-minute walking distance in healthy middle-aged and older subjects. *The European respiratory journal*. 2000; 15: 1132-3.
38. Troosters T, Gosselink R and Decramer M. Six minute walking distance in healthy elderly subjects. *The European respiratory journal*. 1999; 14: 270-4.
39. EuroQol--a new facility for the measurement of health-related quality of life. *Health policy (Amsterdam, Netherlands)*. 1990; 16: 199-208.
40. Schaffler N, Schonberg P, Stephan J, Stellmann JP, Gold SM and Heesen C. Comparison of patient-reported outcome measures in multiple sclerosis. *Acta neurologica Scandinavica*. 2013; 128: 114-21.
41. Mills RJ, Calabresi M, Tennant A and Young CA. Perceived changes and minimum clinically important difference of the Neurological Fatigue Index for multiple sclerosis (NFI-MS). *Multiple sclerosis*. 2013; 19: 502-5.
42. Penner IK, Raselli C, Stocklin M, Opwis K, Kappos L and Calabrese P. The Fatigue Scale for Motor and Cognitive Functions (FSMC): validation of a new instrument to assess multiple sclerosis-related fatigue. *Multiple sclerosis*. 2009; 15: 1509-17.
43. Derksen A, Mokkink LB, Rietberg MB, Knol DL, Ostelo RW and Uitdehaag BM. Validation of a Dutch version of the Neurological Fatigue Index (NFI-MS) for patients with multiple sclerosis in the Netherlands. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. 2013; 22: 2435-41.
44. Bansi J, Bloch W, Gamper U, Riedel S and Kesselring J. Endurance training in MS: short-term immune responses and their relation to cardiorespiratory fitness, health-related quality of life, and fatigue. *Journal of neurology*. 2013; 260: 2993-3001.

Table 1 – Factors considered when discussing/making recommendations regarding outcome measures

Factors considered
<p>Psychometric properties of the outcome measure including reliability, validity, responsiveness to change</p> <p>Self report or performance measure</p> <p>Generic or MS specific measure</p> <p>The time taken to complete the assessment</p> <p>The patient/clinician/ administrator burden involved</p> <p>Any resources required including equipment, specialist skills, space</p> <p>The appropriateness of the measure across the disability range</p> <p>The specificity of the measure in terms of the type of exercise e.g. aerobic, resistance</p> <p>Whether measures were under licence</p> <p>The availability of normative data</p> <p>Time frame e.g. questionnaires relating to previous 7 days or 4 weeks</p>

Table 2 –Stage 1 - each of the 93 categories of the ICF Core set deemed likely/potentially likely or unlikely to be affected by exercise

[illegible]

	Toileting Dressing Eating Drinking Looking after ones health Acquisition of goods and services Preparing meals Doing housework Caring for household objects Remunerative employment Community life Recreation and leisure (n= 27)	(n=11)	(n=15)
--	---	--------	--------

Table 3 Recommended Core Outcome Measures for exercise studies in MS

Dimension	Outcome Measure(s)
Quality of Life	Multiple Sclerosis Impact Scale (MSIS-29) ¹⁴ or MSQoL54 ¹⁵
Energy and drive (fatigue)	Modified Fatigue Impact Scale (MSIF) ¹⁶ (multi-dimensional) or Fatigue Severity Scale (FSS) ¹⁷ (uni-dimensional)
Exercise Tolerance	6 Minute Walk Test (6MWT) ¹⁸
Muscle function/moving around	Timed up and Go (TUG) ¹⁹
Body Structure (body composition)	Waist-hip ratio (WHR) Body Mass Index (BMI) ²⁰

Table 4 Detail of the Recommended Core Outcome Measures

Factors considered	MSIS-29	MSQoL54	MFIS	FSS	6MWT	TUG	BMI	WHR
Psychometric properties of the outcome measure including reliability, responsiveness to change	<p><u>Internal Consistency:</u> Cronbach's alpha values .96 (physical scale), .91 (psychological scale) n=703¹⁴</p> <p><u>Inter-rater reliability</u> ICC 0.58 for physical and 0.34 for psychological²¹</p> <p><u>Test-retest reliability:</u> ICC 0.94 (physical), 0.87(psychological) n=128¹⁴</p> <p><u>ICC 0.87(physical), 0.83 (psychological) n=30²¹</u></p> <p><u>MDC/responsiveness</u> Change in 8 points is clinically significant (baseline to 4 yrs), n= 214²²</p>	<p><u>Internal Consistency</u> Cronbach's alpha values .75-.96 for 12 subscales¹⁵.</p> <p><u>No inter/intra rater reliability data</u></p> <p><u>Test-retest reliability</u> ICC 0.66 to 0.96¹⁵</p> <p><u>MDC/responsiveness</u> No MDC or MICD values. More sensitive to change than generic QOL measures in patients receiving steroid treatment²³.</p>	<p><u>Internal Consistency:</u> Cronbach's alpha values > .94 for total score, physical and cognitive subscores n= 1721 pwMS²⁴</p> <p><u>Test-retest reliability:</u> ICC = 0.85, physical 0.73, cognitive 0.88, psychosocial 0.81 n= 43 ambulatory MS²⁵</p> <p><u>ICC=0.86, physical 0.86, cognitive 0.84, psychosocial 0.78 n= 82²⁶</u></p> <p><u>MDC/Responsiveness</u> MDC: Total score 19.3%, physical 24.7%, cognitive 20%, psychosocial 28.8% in pwMS²⁷</p> <p>Smallest detectable difference (points change): total 16.2, physical 8.9, cognitive 8.0, psychosocial 2.3 in pwMS²⁷</p> <p>SEM 7.3 points, CV 21.4%, MDC: 20.2</p>	<p><u>Internal Consistency:</u> Cronbach's alpha value .93 in 1721 people with MS²⁴</p> <p><u>Test-retest reliability</u> ICC 0.76 n= 43 ambulatory pwMS²⁷</p> <p><u>ICC total score 0.75 in 82 pwMS over 6 month retest period²⁶</u></p> <p><u>MDC/ Responsiveness</u> MDC 20.7%²⁷</p> <p>SEM 0.7 points, CV 10.3%, MDC 1.9 points in 82 people with MS over 6 month retest period²⁶</p>	<p><u>Intra-rater reliability</u> ICC = 0.91; n=40, EDSS range 0 – 6.5¹⁸</p> <p><u>ICC = 0.98, n=12 EDSS 0-6.5²⁸</u></p> <p><u>Inter-rater reliability</u> ICC = 0.94, n= 40, EDSS range 0 – 6.5¹⁸,</p> <p><u>ICC = 0.93 n=19, EDSS ≤6.5²⁹.</u></p> <p><u>ICC=0.96, n=24 EDSS 5-6.5³⁰</u></p> <p><u>Test-retest reliability</u> ICC = 0.96, n=12, EDSS 2.0 – 6.5³¹</p> <p><u>ICC = 0.96 , n=19, EDSS ≤ 6.5²⁹.</u></p> <p><u>ICC=0.959 n=82 EDSS 0-6.5²⁶</u></p> <p><u>MDC/ Responsiveness</u> ±92.16m , n=120, EDSS 0-6.5³²</p>	<p><u>Intra-rater reliability</u> No MS specific data</p> <p><u>Test- retest reliability</u> ICC=0.91 (0.83-0.95), EDSS <= 4: 0.84 (0.66-0.93), EDSS > 4: 0.88 (0.76-0.95)³³</p> <p><u>ICC=0.97, n=24 EDSS 5-6.5³⁰</u></p> <p><u>MDC/ Responsiveness</u> MDC 12.6s n=24 EDSS 5-6.5³⁰</p>	<p>Each laboratory to determine its own reliability data.</p> <p>Body mass index is calculated by dividing body weight in kilograms (kg) by height (m) squared (kg/m²).²⁰</p>	<p>Each laboratory to determine its own reliability data.</p> <p>Hip circumference is measured at the widest portion of the buttocks and waist circumference is measured at the narrowest part of the torso between the iliac crest. In obese people, waist circumference is measured on a horizontal plane around the abdomen at the level of the iliac crest.²⁰</p>

			points n= 82 ²⁶					
Self report or performance measure	Self Report	Self Report	Self Report	Self Report	Performance	Performance	Objective measure	Objective measure
Generic or MS specific measure	MS Specific	MS Specific	MS Specific	Generic	Generic	Generic	Generic	Generic
The time taken to complete	10-15 minutes	10-20minutes	5-10minutes	5 minutes	6 minutes	1-2minutes	5 minutes	5 minutes
The patient/clinician/ administrator burden involved	Minimal	Minimal but scoring is time consuming. (10-20 minutes)	Minimal	Minimal	Minimal	Minimal	Minimal	Minimal
Any resources required; equipment, specialist skills, space	None	None	None	None	Stopwatch, Two small cones to mark turning point, A chair that can be moved anywhere on the track, Worksheets & clipboard,	Stopwatch, 47-cm-high chair with arm and back supports, cone, tape, 3m walkway	Scales to weigh the person and a stadiometer for height measurement (or measuring tape)	Measuring tape
The appropriateness of the measure across disability range	Applicable across the disability range	Applicable across the disability range	0.0 – 7.5 ¹² . May not be appropriate for EDSS score 8 and above	0.0 – 7.5 ¹² . May not be appropriate for EDSS score 8 and above	Ambulatory people only, rests are permitted	All ambulatory people with MS can complete	Applicable across the disability range	Applicable across the disability range
The specificity of the measure in terms of the type of exercise e.g. aerobic, resistance	Not specific to type of exercise as captures information on quality of life	Not specific to type of exercise as captures information on quality of life	Not specific to type of exercise as measures the effects of fatigue on physical, cognitive and psychosocial	Not specific to type of exercise as measures severity of fatigue and it affect on activities and	Not specific to type of exercise as measures walking endurance	Not specific to type of exercise as captures sit to stand, balance and gait functions	Not specific to type of exercise as captures information on body	Not specific to type of exercise as captures information on body

			functioning	lifestyle			composition	composition
Licence required	No	No	No (under copyright)	No	No	No	No	No
The availability of normative data	Minimal change in physical score (0.9) in 553 pwMS over 300 days ³⁴	No	No	Limited ³⁵	Yes ³⁶⁻³⁸	Yes ³⁶	Cut-points for health ²⁰	Cut-points for health ²⁰
Time frame e.g. questionnaires relating to previous 7 days or 4 weeks	2 weeks	4 weeks	4 weeks		Current performance	Current performance	Data relates to measures taken at that time point	Data relates to measures taken at that time point