Cronfa - Swansea University Open Access Repository

Outcome measures for Multiple Sclerosis: A review

Lisa A. Osborne[1], J. Gareth Noble[1], Inocencio D.C. Maramba[1], Kerina H. Jones[1], Rodden M. Middleton[1], Ronan A. Lyons[1], David V. Ford[1], & Phil Reed[2]

[1] College of Medicine, Swansea University, U.K.

[2] Department of Psychology, Swansea University, U.K.

Short Title: Outcome measures for MS.

Abstract

Background: This review determined the most commonly used, and reliable, measures for assessing clinical outcomes for Multiple Sclerosis (MS).

Objectives: It was anticipated that this would facilitate the development of a common set of metrics, and aid reaching a consensus regarding the outcome measures that are typically used in the field of MS clinical research.

Major Findings: A thorough literature review of clinical outcome measures for MS produced 166 measures that have been used in this context. This list was then refined by discussion with a panel of consultant neurologists, which reduced the list to 23 commonly employed tools. This shortlist was then further refined through surveying 41 centres for MS treatment, which reduced the shortlist to 16 measures. The properties of these scales, in terms of their symptom/function domains, their specificity for MS, their administration characteristics, and their reliability and validity for MS, are all discussed.

Conclusions: Conclusions regarding the development of potential sets of assessment measures for MS, that encompasses broad symptom/function domains, and which are sensitive to the practical requirements of administration within clinical contexts, are explored.

Introduction

Multiple Sclerosis (MS) is a chronic condition of the central nervous system[1,2]. Although a central database for MS in the U.K. is being piloted[3-5], there are currently no definitive figures regarding those it affects. Estimates suggest that anywhere between 70,000 and 100,000 people in the U.K. have MS[6-8]. MS is most often diagnosed between the ages of 20 and 40 years, although it can be diagnosed earlier or later in life[3,6], and it appears to be one of the most common disabling neurological diseases among young adults[9]. MS is a complex condition that presents a wide spectrum of severities, symptoms, and impacts on functioning; the neurological damage impacts on physical[10-12], cognitive[13-15], and psychological and emotional[16-18] functioning, as well as on quality of life[19-21]. In order to enhance medical practice in diagnosis, care, and treatment, it is important to determine which measures are best employed in the assessment of MS symptoms/functions and clinical outcomes[22,23]. This topic has recently been the subject of some study and debate in the context of clinical outcomes[22,23], and any databases or and registries aiming to further knowledge about MS clearly require strong measures at their cores[23]. In fact, recently, some such databases have been criticised on the basis of the measures selected[24]. The questions, raised by numerous task forces concerned with MS over the last two decades about the best set of measures for clinical outcomes for MS that can detect changes in impairment and evaluate a person's MS symptoms and functioning, remain largely unanswered[22,25].

There are numerous measures available for the study of MS[26], and a number of reviews of these assessment instruments have previously been conducted[20,22,27-34]. These reviews highlight several commonly used instruments, as well as outlining their properties. However, these reviews are not

always comprehensive in their coverage of the literature, frequently focusing on only a few selected measures. They do not always address the full range of MS symptomatology or all aspects of a person's functioning, focussing instead on only one specific aspect of MS[28,32,34,35]. Previous work does not always attempt to define standard and comprehensive packages of measures to encompass all aspects of MS (i.e. physical, cognitive, psychological and emotional, and quality of life functions), although this is increasingly recognised as an important step[36,37]. Moreover, with a few exceptions[38], often previous review articles do not account for typical clinical-care usage, but rather are inclined to be research-oriented in outlook[12,32].

The review that is reported here aims to provide an up-to-date, consensus-based overview of the most extensively utilised clinical outcome measures for MS that could detect changes in levels of impairment, compatible with the practicalities and demands of clinical care, detailing their relative strengths, and their focus on particular symptom/function domains. In doing so, this review addresses current practice and views concerning the most effective instruments currently employed for assessing MS symptoms/functionality and treatments. This review has multiple objectives: (1) to survey and critique the measures in common use; (2) to examine the strengths and weaknesses of these measures in the context of MS data collection (e.g., their reliability and validity for this population); and (3) in the light of these findings to offer a helpful resource for clinicians and researchers working in the field of MS to enable them to choose the set of measures that are best suited for their particular needs. The study's outputs will aim to be a helpful resource for clinicians and researchers working in the field of MS.

Review Method

Search Strategy

Search engines and computerised bibliographic databases were deployed to identify papers that have used measures of clinical outcomes for MS (PubMed, PsychINFO, Google Scholar, Web of Knowledge). In all cases, the search phrases, "clinical-outcome-measures" and "multiple-sclerosis", were employed together, using a backward chronological saturation search strategy. Such a search strategy moves from the most recent year (2012) backwards, in chronological order, through the published literature, year by year. For every article identified, its methodology was examined to determine which measures had been used. This process continued until no further measures were identified from the search. The 'saturation principle' was used to determine when to stop (i.e. when no new measures were coming to light).

This method was designed to produce a list of outcome measures that reflects current and recent usage. This is important, given the fast-moving nature of medical and health care developments as a whole, and in this field, in particular, with many instruments rapidly superseded and replaced by alternative measures, or are altered and updated with newer versions.

Validation

To ensure that there were no major omissions from the initial list of outcome measures, two additional search and validation strategies were adopted. Firstly, the full list of instruments, derived from the current search, was compared to those listed in a number of recently published reviews of the literature on MS outcome measures[20,22,27-30,32-34]. Secondly, the full set of measures derived was compared to the list of measures provided by the NICE Guidelines[26]. This

rigorous search and validation process identified an initial total of 166 instruments that have been employed extensively as MS outcome measures.

Refinements

This extensive list was subjected to two stages of refinement, in order to produce a summary of contemporary clinical outcome measures most often used for MS. Firstly, the full list was provided to an independent panel of consultant neurologists with expertise in MS, convened by the U.K. MS Society, for scrutiny and comment. This process involved discussion of these measures, moderated by a member of the MS Society who was independent to the authors of this study. The panel members were asked to comment on the full list, and to indicate which measures they typically employed for the assessment of MS outcomes. This refinement process was directed at developing a much reduced version of this list. This first refinement process resulted in 23 measures being identified.

The 23 measures identified as most used in MS clinical contexts, were sent, as part of a questionnaire, to 82 U.K. centres involved in the treatment of MS. These centres were asked to indicate which of these measures they employed, and whether they used any other measures in addition to those listed. The teams who were consulted at each of the treatment sites were interdisciplinary, and typically comprised a Consultant Neurologist, a Head MS Nurse, a Physiotherapist or a Speech Therapist or an Occupational Therapist. The teams all contributed to the questionnaire responses, and the Head MS Nurse usually returned the questionnaire after consultation with their other team members. Responses were obtained from 41 centres (50%).

These responses indicated that none of the centres used measures that were not already included in the shortened list from the expert panel consultation. All of the 23 measures, except for two (Brief Repeatable Battery of Neuropsychological Tests, and the Percentage of Patients Remaining Relapse Free) were employed by at least one of these centres. There were 21 measures identified as being in contemporary use for MS clinical outcome assessments. The mean use across the centres for any specific scale was 20% of the centres ($\pm$ 17%), with the use ranging from 2% of the centres employing the least-used measure, to 76% of the centres employing the most-used measure (see Figure 1).

Figure 1

Measurement Scales

The identified measures were subject to further analysis by searching the literature in order to identify their features, and their psychometric properties. The standard types of psychometric properties were identified (see Figure 2) to enable the measures to be compared with one another. In all cases, at least two independent assessments of these properties were identified from the literature, where possible, from sources other than the author. Of the remaining 21 measures, four were removed. Clinical Relapse Rate/Severity/Time to First Relapse was removed, because it could not be considered as a measure with psychometric properties; Quality-Adjusted Life Years, depends on other measures for its usefulness; and a further two, Evoked Response Potential, and Magnetic Resonance Imaging, were generic technologies that encompass a wide range of measures, which are not readily susceptible to psychometric analysis.

Additionally, details regarding the Fatigue Impact Scale, and its modified version, were merged into one for the purposes of this review.

Figure 2

Figure 1 shows the measures categorised by the predominant domains that they address. In addition, each instrument is classified as MS-specific or generic, and as clinician-administered or patient-completed. Figure 3 shows the psychometric properties, as measured on MS samples, and main qualities of each measure.

Figure 3

Physical Disability

The Modified Ashworth Scale[39,40] measures muscle spasticity for each joint on a 6-point scale. The resulting score has variable inter-rater agreement, some being good (86%[39]), but this reliability is better for upper joints (80.6%) than for lower joints (63.9%)[41,42]. As motor function can vary dramatically between extremities within individuals with MS, it makes little sense to evaluate the internal consistency of this scale, although some attempted this, and found moderate values (0.78[37]). The test has varying test-retest reliabilities, 93.4% for upper joints, and 71.1% for lower joints, with some studies reporting very low test-retest reliability[37]. However, the variability of these reliabilities across joints has caused some concern[43,44]. The resulting spasticity score has rather poor concurrent validity with similar disability scales, such as the

Spasm Frequency Scale[45], and electromyography[46], and a poor correlation with the EDSS[37].

Upper joints have better concurrent validity than lower joint scores[47]. The variable reliability and

poor validity have lead some to question its usage for MS[48], and the modified modified version

has been produced, which appears to show better psychometric properties in its initial trials[49].

Expanded Disability Status Scale (EDSS)[50] is a measure of impairment and impacts of MS,

which scores disability on a 20-point scale (0 = normal to 10 = death), on the basis of a

neurological examination. In addition to the overall score, there are also 8 associated functional

systems assessment scales. There is a patient-rated equivalent, the *Patient Determined Disease*

Steps, which is a single-item scale, relating to 8 levels of disability. This is a generic measure of

disability, but it is mainly used for MS, and it correlates between 0·70[51] and 0·93[52] with the

clinician-administered EDSS. The EDSS has moderately good reliability characteristics, having

inter-rater correlations measured as between 0·32 and 0·76[12], and 0·5 and 0·7[53]. It has an internal

reliability of 0·88 to 0·96[54], and test-retest reliability reported as between 0·42 and 0·66[12], and

0·61 and 0·94[55]. It also has good concurrent validity, in that it correlates well with the Multiple

Sclerosis Functional Composite score, with measures of disability (0·84)[55], and with physical, but

not mental, measures of quality of life[12,55,56]. It correlates well with patient-rated measures of

physical disability[51]. It has poor sensitivity (24%), but good specificity (79%)[57]. However, a

major drawback is its lack of responsiveness to clinical change (effect size = 0·1)[12,58], being

better for less severely disabled individuals[59]. In particular, changes in the EDSS score from one

time to another often fall within the error that would be typically expected on the basis of inter-

rater agreement[60]. To help overcome these problems it has been suggested that two EDSS

determinations are made prior to intervention, and either the mean value (to increase power) or

the lower value (to increase sensitivity) be employed[61]. It is generally agreed that it would need substantial modification to be used as a main tool of clinical change[62], but its widespread use and long history mean that some have suggested that it continue to be used as part of a battery of assessment tools[22].

Functional Independence Measure (FIM)[63] assesses the ability of a patient to conduct routine daily activities, also often used for older people and those who have suffered a stroke. The instrument contains 18 items, which produce a composite measure of independence when assessed: 13 of these items are concerned with motor function (based on the Barthal Index); and 7 items are related to cognitive function. A shorter version, the alpha-FIM, is available, which has only 6 items[64]. All items are rated on a 1 (total assistance required) to 7 (full independence) scale. It can be combined with the Functional Assessment Measure (FAM), which adds 12 items concerned with cognition, similarly completed by a health care professional[65]. The FIM has an inter-rater reliability of between 0.83 and 0.99[66-68], although this is greater for the motor items (0.95) than for the cognition items (0.78)[67], suggesting that the FIM should not be used without the FAM for assessment of cognitive independence[69]. The FIM also has good internal reliability (0.94[66], 0.98[12]). It has good test-retest reliability, measured as between 0.95[67] and 0.99[12]. It has good concurrent validity with both the EDSS[66], and the Barthal Index[70]. As the FIM was heavily based on the Barthal Index for motor items, this is not surprising. However, the validity is less good for the neurological aspects that it measures. There are no data for its sensitivity and specificity. A key criticism is that it is not responsive to change[12], with effect sizes of between 0.3[69] and 0.46[12], and lower for the cognitive items. This lack of responsiveness to change may be the result of floor and ceiling effects[71].

Guy's Neurological Disability Scale (GNDS)[72] is a measure of general clinical disability, and asks the patient to assess their current clinical state over the last month (a patient self-administered version is also available[73]. It gives an overall measure of disability and sub-scores covering 12 domains. It has an internal reliability of around 0.80[72,74]. This instrument has a very good test-retest reliability of over 0.96[72,75] for the overall score, and of 0.68 to 0.99 for the sub-scales. However, it has mixed results in terms of concurrent validity; correlating 0.64 with the EDSS, and 0.76 with the Barthal Index[75]. As would be predicted, it has lower correlations with scales that do not measure physical disability[76]. There are no data on its sensitivity or specificity. The GNDS appears to be reasonably responsive (effect size, 0.58)[12], although lower effect sizes have also been noted[77]. Its duration of administration, and its uncertain responsiveness, means it tends to be used largely in clinical practice, rather than for trial purposes[27].

Multiple Sclerosis Functional Composite (MSFC)[78] assesses the level of impairment across three domains at the time of testing (i.e. leg function is assessed by a 25-feet walking task, arm function by the nine-hole peg test, and cognitive function by use of the paced auditory serial addition test). The assessment provides analysis of the above three aspects of functioning, as well as an overall composite measure (although this summary measure has been critics as being 'abstract' and difficult to interpret clinically[22]. The measure has an inter-rater reliability of 0.95[79] to 0.96[80] for the composite score, and 0.93 to 0.99 for the components[80]. It has an internal reliability of 0.97 for the overall composite score[81]. The test-retest measures are good, ranging between 0.87[56] and 0.90[82] for the overall score, and 0.93 to 0.98 for the subtests[80]. However, there

are some practice effects that are particularly noticeable on the first few applications[82]. It has

good concurrent validity scores with a range of other measures of disability: 0.80 with the

EDSS[83,84]; and it correlates well with MRI measures of brain atrophy[79]. It shows weaker

association with measures of emotional functioning[79]. It has moderate sensitivity (51%)[79], and it

is moderately good at predicting both disability[85] and MRI results[86]. It has a strong (86%)

specificity[79]. Its responsiveness is measured as being between 0.62 and 0.71 in terms of the area

under the curve in the ROC[87], and has a reported effect size of 0.50, although responsiveness to

clinical trails can be compromised by the absence of a test for vision[22], and this value is better for

the leg function than for the upper limb and cognitive measures[56]. Moreover, the responsiveness

of the cognitive component is not strong[88], and this should not be relied upon as a sole measure

of cognitive change for MS. As a result many have suggested replacing the existing cognitive

component (the paced auditory serial addition task), with the symbol digit modalities test, to

enhance the test[22].

Multiple Sclerosis Walking Scale-12 (MSWS-12)[77] assesses patient-reported disruption to their

walking, and its 12 items measure patient-perception of their walking quality during the past 2

weeks. This measure has a strong internal reliability of 0.97[89-91], and good test-retest reliability

ranging from 0.86[91] to 0.96[89], and over 0.78 for the individual items[77]. It has good concurrent

validity, as it correlates 0.80 overall with the EDSS (although this relationship is mainly driven

by strong correlations at lower levels of disability), and 0.77 with the MSIS-29 physical scale[90],

and 0.82 with physical scales from the SF-36[89]. However, its correlation with accelerometry

scores are moderate, and range from 0.38[92] to 0.70[93], as are its correlations with other objective

measures of walking[94]. There are no data on its sensitivity and specificity. A strong feature of

the tool is that it has good levels of responsiveness[35], with effect sizes being greater than 1·00[77,95], but the variance between differently disabled groups may make it less than optimal for comparing between samples[96].

Cognitive Impairment

Mini-Mental State Examination (MMSE)[97] screens for dementia and cognitive impairment at the time of testing. It comprises 11 items that measure a number of domains (a standardised version exists, but this has not been tested for MS[98]). It has a good internal reliability of 0·89 to 0·95[97], but variable test-retest reliability, ranging between 0·65 and 0·89[97]. It does display some moderate concurrent validity – it correlates 0·78 with the verbal IQ score from the WAIS[97]. However, it has been noted that it is not well suited for the patterns of disability seen in MS[99], and it has poor sensitivity for this population of between 28%[100] and 36%[101]. It does have good specificity (89%)[101], but poor, and unreliable, responsiveness, given its weak test-retest reliability[102]. Given the latter problems, some have recommended that it not be used for people with MS[103], and it is more commonly employed as part of a battery of tests to describe the sample characteristics, rather than to assess change due to an intervention[104].

Psychological/Emotional

Beck's Depression Inventory (BDI-II)[105,106] measures depression over the last week. There are 21 items, each rated on a 4-point scale, relating to symptoms of depression, which gives an overall score. It has an internal reliability score for people with MS of between 0·86[107] and 0·94[108]. It has a test-retest reliability of 0·93 across people with a range of disorders[105]. It has good concurrent validity for MS[109], and correlates well (0·65) with psychiatric ratings of

depression, and with the SF-20 sub-scales[108], and 0.71 with the Hamilton Depression Scale[105]. It

has between 81%[110] and 84%[111] sensitivity, and 70% specificity, in all phases of the disease[111].

Its responsiveness is also reasonable[109,112]. However, there is some discussion about whether

some of the items from the BDI-II should be removed for use with MS, as they overlap with

some of the symptoms and impacts of MS[107,113].

General Health Questionnaire (GHQ)[114] is used for assessing general psychiatric problems over

the prior 'several weeks', and it comes in many versions, including: the GHQ-12 (which gives an

overall score); the GHQ-28 (which gives scores for somatic problems, anxiety, social

dysfunction, and severe depression, as well as an overall score); the GHQ-30 (which gives an

overall score); and the GHQ-60 (which gives an overall score). The most used for MS are the

GHQ-12 and GHQ-28, being the shortest versions. The GHQ-12 has an internal reliability of

0.91[115], and the GHQ-28 has an internal reliability of 0.90[116]. The GHQ-28 has a test-retest

reliability of 0.69 in a general population[116]. There are reports of a retest/practice effect

impacting on the score[117]. The GHQ-28 has a moderate concurrent validity (0.44 to 0.66) with a

range of health measures[76], and correlates 0.83 with the present state examination[118]. There are

mixed reports about the sensitivity of the GHQ-12, these being 36% to 67%[76] or 72% to 92%[119],

the GHQ-28 being reported to have better sensitivity[120], and is recommended as one of the stem

tools for identification of depression in chronic illness[121]. The GHQ-12 has good specificity,

being measured as over 74%[76,119,120]. Reports on the responsiveness of the GHQ-12 are also

mixed, varying from effects sizes of 0.15[122] to 0.51[123].

Hospital Anxiety and Depression Scales (HADS)[124] assess levels of anxiety and depression, excluding somatic symptoms to avoid overlap with any physical symptoms of the patient. Originally designed for use by hospital general medical outpatients, it has since been widely employed in primary care settings, and also in on-line surveys[125], and contains 14 items (7 for anxiety and 7 for depression), that relate to the last week. Patients are recommended not to take very long to respond, as their immediate replies should better reflect their actual state than more considered reactions[124]. This instrument has excellent reliability, having an internal reliability of 0.83 for anxiety[21,126], and of 0.81[126] to 0.82[21] for depression. Its test-retest reliability is 0.89 for anxiety and 0.86 for depression[127]. It has reasonable concurrent validity correlates between 0.49 and 0.83 with similar instruments[128], and 0.62 for anxiety, and 0.55 for depression, with the appropriate Beck scales[76]. It has mixed reports concerning sensitivity, it has 88.5% sensitivity for anxiety, and a 90% sensitivity for depression, when tested against objectively rated symptoms[129], but between 29% and 46% for anxiety, and 25% and 46% for depression, when measured against other self report scales[76]. Its specificity is excellent, being measured as between 80% and 90% for anxiety, and 84% to 90% for depression, regardless of the anchor measure used[76,129]. However, the anxiety scale tends to be more sensitive to General Anxiety Disorder than to other forms of anxiety[129].

Quality of Life

EuroQol-5 Dimensional Questionnaire (EQ-5D)[130] is a measure of the impacts of disease on various aspects of health. There are 5 specific, Likert-type, questions addressing the patient's state on the day of assessment (plus optional questions relating to socio-economic variables). It has reasonable test-retest reliability reported as 0.63-0.80[131], and 0.81[132]. It has reasonable

concurrent validity, correlating 0.70 with the SF-60[132], and 0.55-0.81 with a range of similar generic quality of life instruments[12]. However, its relationship with EDSS scores in MS are uncertain[133], and it does not correlate with objective measures of health and functioning[134]. It is not regarded to have sensitivity for MS[76], and the absence of a fatigue score is an issue in this regard[135], and there are no direct specificity data. Its responsiveness is less than that of other instruments[12], with some ceiling effects and with 40% of severely disabled individuals likely to omit the items relating to physical status[132]. This measure is considered good for population description, but it may be too insensitive for measuring individual changes. Although it is commonly used for cost estimates in health-economics[136], some have suggested that disease-specific tools, such as the MSIS-29, are more sensitive in this context[137].

Fatigue Impact Scale and Modified (FIS/MFIS)[138,139] measures the impact of fatigue on daily life over the last 4 weeks. This instrument comprises 40 items, and gives an overall fatigue impact score and three sub-scales. (The 21-item MFIS[139] is a reduced version of the FIS, taking 5 to 10 minutes to complete, giving an overall score and three subscales; an even briefer 5-item version takes 2 to 3 minutes to complete). The FIS has good indicators of overall internal reliability of 0.98[138], and 0.85[140], with the individual sub-scales all having internal reliabilities of over 0.87. The overall MFIS score has an internal reliability measured from 0.81[139] to 0.92[141], with sub-scales raging between 0.88 and 0.92[141]. The test-retest reliabilities of the overall FIS score is measured as between 0.81[142] and 0.93[122], with the sub-scales ranging from 0.68 to 0.85 [138,142]. For the MFIS, the overall test-retest reliability varies between 0.82-0.85[141,143,144]. There is moderate concurrent validity of the FIS with the SF-36[142], and between the MFIS and the Fatigue Severity Scale (0.66)[144], 0.68 [145]. In terms of its sensitivity, 78% of people with MS are identified

correctly by the FIS[138], with the MFIS have a greater then 90% sensitivity[146]. There are no data

on specificity for the FIS, but the MFIS has greater than 90% specificity[146]. The FIS has a greater

than 0.7 effect size in terms of its responsiveness[147], with the MFIS having effect sizes of greater

than 0.5[148,149], but it is not greatly responsive over short periods of time[144]. Recently, the

usefulness of the overall fatigue impact score of these scales has been called into question,

although the sub-scales, used on an individual basis, are considered to be safe[150], and, as a

result, the 22-item Unitary FIS (U-FIS) has been developed to give an overall fatigue score[151].

Nevertheless, this is the tool recommended to assess fatigue-related quality of life[152].

Leeds Multiple Sclerosis Quality of Life Scale (LMSQoL)[153] is a measure of the impact of MS

on quality of life at the time of completion. This scale has 8 items, each scored on a 4-point

Likert-type scale, and it gives an overall index of quality of life. The scale has a good internal

reliability, being reported as between 0.71 and 0.86[21,76,153], and good test-retest reliability of

0.85[153]. While there are some reports of moderate concurrent validity: 0.68 with the SF-36, and

0.83 with measures of well-being[153], and it correlates with diary reports of the impact of MS[154],

recent questions have been raised about its validity when compared to the MSQoL[76]. There are

no data for its sensitivity and specificity. It has moderate responsiveness, having reported effect

sizes of 0.34[76,155], and 0.45[156], with little sign of floor or ceiling effects[157].

Multiple Sclerosis Impact Scale (MSIS-29)[158] measures the impact of MS on daily living over

the preceding 2 weeks, rather than on the resulting quality of life. Its 29 items produce an overall

score and sub-scales relating to both physical and psychological domains. The overall score is

considered to be of debatable use[159]. The internal reliability of the two sub-scales is very good

ranging from 0.88 to 0.96 for the Physical, and 0.85 to 0.91 for the Psychological[158,160-162], as are

their test-retest reliabilities, which range from 0.86 to 0.94 for the Physical, and from 0.81 to 0.87

for the Psychological[158,161-163]. The score correlates moderately well with similar instruments,

giving this scale reasonable concurrent validity: the Physical scale correlates 0.66 with the EDSS

and 0.69 with the GNDS[162], and greater than 0.5 with other measures of disability[160,164]; the

Psychological scale correlates greater than 0.6 with other scales measuring mental functioning[160].

The sensitivity is good[137]: 78% for the Physical scale[161,165]; and 73% for the Psychological

scale[161]. Reports on its specificity are mixed, some results being good (greater than 80% for both

scales[161], but some being only moderate (51% for the Physical scale[165]. The responsiveness of

the Physical scale, which is good (0.82 effect size), is better than that of the Psychological scale,

which is moderate (0.66 effect size)[122]. The area under the curve in its ROC is good when

anchored against the EDSS, 0.72[165], but lower values have been reported[166].

Multiple Sclerosis Quality of Life-54 (MSQoL-54)[157] assesses the quality of life of individuals

with MS over the last 4 weeks (the 54 items include the 36 items from the generic SF-36

questionnaire). This tool gives an overall score for quality of life, comprising two health sub-

domains (Physical and Mental). The overall score has an internal reliability measured as between

0.84[167] and 0.96[168], with the sub-scales having internal reliabilities ranging from 0.75 to 0.96[157].

The test-retest reliability coefficients range between moderate 0.61 and good 0.96[157,169]. The

overall measure correlates moderately with MS symptoms measured by the ICD[167], it also

correlates moderately with the EDSS and MSFC (0.49 to 0.67), and with the Leeds MS QoL

scale[170]. Both the Physical and Mental Health sub-scales correlate well with the Fatigue Severity

Scale[171]. There are no data on its sensitivity and specificity, although the addition of 18 MS-

specific questions to the generic SF-36 should improve these properties relative to the generic form[157]. Reports on its responsiveness are mixed[172] with effect sizes of over 0.7 for the physical scales, and between 0.57 and 0.7 for the mental scales[166], while other reports on responsiveness have been poor[123]. There has been some criticism of potential floor and ceiling issues that may limit its usefulness as an outcome measure[28].

Short Form SF-36 Health Scale[173] measures the impact of a disorder on the functioning of an individual over the last 4 weeks. Derived from the GHQ-28[114], this form contains 36 items, producing an overall score relating to impacts on health, as well as two sub-scales (physical health and mental health). The overall scale has an internal reliability of 0.67 to 0.94[173], with the sub-scales ranging between 0.77 and 0.96[54,174]. The test-retest score results are mixed, and are between 0.64 and 0.96 for the subscales[174]. However, there have been recent suggestions that the psychometric properties of the two main scales (physical and mental) are not as good as the overall score, or those of the sub-scales on which they are based[158]. It has moderate concurrent validity (0.6) with symptom severity[174]. The physical scale has moderate correlations (0.6) with the FIM and the EDSS, and the mental sub-scale correlates 0.5 with the GHQ[123,133]. It has moderate sensitivity for body functions and activities in MS[175], but there are no data on its specificity. The effect sizes show poor responsiveness, ranging from 0.01 to 0.30[123,176].

Discussion

The findings suggest a relatively wide range of instruments are commonly used in MS assessment, and those employed measure many varied MS symptoms/functions, and reflect the need to capture the diverse nature of this disorder's symptomatology and impact on a person's

functioning. These scales evaluate the symptoms/functioning in each of four broad categories: physical (e.g., physiological, fatigue, movement); cognitive (e.g., memory, attention); psychological and emotional (e.g., depression, anxiety); and quality of life (i.e. impacts of the symptoms/functioning on daily life and living). Naturally, there is overlap across the aspects that an individual scale measures, and between these broad categories. It is not always straightforward to ascertain which domain a questionnaire indexes; the content of some measures covers more than one domain. For instance, there can be difficulty discriminating the direct symptoms of MS and their effects on functioning from the impacts of those symptoms/functions on daily life. Some aspects of MS (e.g., gait, pain, incontinence, depression) may be physical or psychological problems and also quality of life problems, and it is not necessarily easy to divide them into separate domains. The impact of physical issues on an individual (e.g., producing depression) is not readily distinguishable from their impacts on what an individual can or cannot do (i.e. their quality of life). In developing test-sets, attention should be paid to these issues. However, it is useful to compare the assessment tools along the four dimensions, to allow comparison between the measures, and to develop appropriate MS assessment sets (see Figures 1 and 3).

In terms of physical symptoms, the measures that address these mainly focus on mobility and motor function, and deal with central nervous system damage. The majority of such tools are clinician-administered (Modified Ashworth Scale, EDSS, FIM, MSFC), with only two such scales being patient-completed (GNDS and MSWS-12). This may reflect the facts that these symptoms/functions are observable, and measurable, by the clinician, and that a patient's judgement of their physical symptoms and functioning may be confounded with other factors

(e.g., mental state). Of the clinician-administered scales, three deal largely with mobility and motor function (Ashworth, EDSS, and MSFC). All of these give reliable measures of the present levels of disability. The patient-completed scales concerning motor symptoms/functioning (some aspects of the GNDS, and the MSWS-12) give reliable measures of current levels of disability, although not as reliable as the clinician-administered scales. However, the clinician-administered scales (especially the EDSS) suffer from a common problem of not being particularly responsive to change in disability status, possibly resulting from the nature of the physical problems themselves, in that, once mobility has been impaired, it may not then readily return to its former unimpaired state (especially in some types of MS). The patient-completed scales appear more responsive to change in disability status, but it is not clear whether this reflects their measure of the disability itself, or a person's perception of their disability, which may be different, and the latter may be responsible for responsiveness to change. In this context, it is informative to consider current models of disability. For example, the World Health Organisation's (WHO) International Classification of Functioning, Disability, and Health (ICF) acknowledges that 'disability' is a constructed experience – that is, the person's own perception of their symptoms/functioning, in addition to any objective assessment of those issues, is a prime aspect of disability. In addition, the literature suggests that better personal perceptions of health are stronger predictors of confidence regarding a person's involvement and participation in their community (which is also a key element defining disability in the WHO's ICF framework).

There are only two clinician-administered measures that deal with cognitive impairments (e.g., memory, attention, problem solving, information processing), which are the MMSE, and one component of the MSFC. However, these measures are rather poor, both in terms of their

psychometric properties, and in their lack of sensitivity to MS cognitive impairments. In addition, neither of these captures the potential range of impairments, focusing only on very specific and limited tests. These criticisms are also potentially true of the patient-completed measure of cognition contained in the GNDS. The measurement of this domain, by the commonly employed scales, is potentially the weakest, and some consideration should be given to employment of better validated tests of cognitive function, or, at least, some of their sub-scales, such as the WAIS.

A third area of symptomatology or functioning that some of these instruments address is the psychological and emotional impacts of MS. Some of these symptoms/functions may result from CNS damage, but many will also result from the impact of MS on the ability of a person to conduct activities in which they wish to engage. Due to the subjective nature of these symptoms and their impacts on functioning, all of the scales that assess these aspects are patient-completed. In terms of depression, both the BDI and the HADS have very good psychometric properties for MS. The HADS has a slight advantage in that it is shorter, and it measures both anxiety and depression. The EQ-5D, GHQ-28, GNDS, and the SF-36, all contain measures of psychological function, that correlate with the BDI and HADS, but they are not necessarily related to clinical definitions of depression, and also reflect the impacts of these problems, rather than the problems themselves. In terms of fatigue, the FIS and MFIS are fatigue-specific instruments, of which the MFIS is thought to have an advantage psychometrically, especially when used for the separate impacts of fatigue on physical and mental functioning. The only measure that deals specifically with pain symptoms is the EQ-5D, although other measures assess the impact of pain on limiting

activities. All of these measures have reasonable psychometric properties, but not all of them are MS-specific, and this may merit consideration.

In regards to quality of life, the psychometrics of almost all of these measures are reasonably robust, although not all of these measures are MS-specific. Typically, it is the case that, the longer the measure, the better the psychometric properties. Most of these longer measures give an index of both physical and mental quality of life functioning. Of the nine quality of life measures (EQ-5D, FIS, GHQ, LMSQoL, MFIS, MSIS-29, MSQoL-54, SF-36), four appear to offer very good coverage of quality of life issues, as well as having excellent psychometric properties (GHQ, MSIS-29, MSQoL-54, SF-36). Of these, the MSIS-29 is MS-specific, the MSQoL-54 has some MS-specific questions, and also encompasses the generic SF-36. However, in choosing an appropriate questionnaire, issues regarding the MS-specificity of a scale's items should be considered. It has been recommended[130] that the MFIS is used, when working with people with MS, as it was developed on an MS population, however, the MSIS-29 is MS-specific, and this measure offers better psychometric properties than the MFIS. The generic GHQ and the SF-36 do similar things to one another, as the SF-36 is derived from the GHQ, and so would not be employed together. The MSQoL-54 is MS-specific, and encompasses the SF-36, and it is similar, in itself, to the MSIS-29, which is shorter. So a choice may be between using the GHQ and the MSIS-29, or using the MSQoL-54. An ideal approach is to choose a set of tests in which there is some overlap across the various measures; for example, when assessing quality of life, it would be desirable to use measures that examine the impact of physical symptoms/functioning in conjunction with other measures that examine the objective physical symptoms/functioning. In addition, a generic HRQoL measure (such as the EQ-5D) should be

included in any set of outcome measures to enable comparison with population norms, with other conditions, and to facilitate health economic analyses using QALYs, which generic HRQoL measures can help to produce.

In summary, the key to better assessment of MS is to develop tailored balanced sets of measures that are both complementary and comprehensive. Given the multiple symptom/function domains of MS, which vary in quantity and quality across patients, it would be very difficult to prescribe a single recommended battery of measures to detect changes in impairment; these would vary according to the needs of the clinician, researcher, or patient. Such assessment sets should include measures for each of the four areas discussed above, with consideration to time efficiency, minimising replication or scale redundancy, and not fatiguing the patient. Provided here is a resource from which suitable sets of commonly-used measures can be selected, determined by the specific requirements of each situation. This review could aid such developments, and be a useful resource in itself.

Acknowledgments

Disclosure Statement

References

1.   McDonald, W.I., Compston, A., Edan, G. et al. Recommended diagnostic criteria for
Multiple Sclerosis: Guidelines from the International Panel on the Diagnosis of Multiple
Sclerosis. Annals of Neurology 2001; 50: 121–127.

2.   Polman, C.H., Reingold, S.C., Edan G. et al. Diagnostic criteria for Multiple Sclerosis: 2005
Revisions to the "McDonald Criteria". Annals of Neurology 2005; 58: 840–846.

3.   Ford, D.V., Jones, K.H., Middleton, R.M.. et al. The feasibility of collecting information
from people with Multiple Sclerosis for the UK MS Register via a web portal: characterising a
cohort of people with MS. BMC Medical Informatics & Decision Making 2012; 12: 73.

4.   Osborne, L.A., Noble, J.G., Lockhart, H., Middleton, R.M., Thompson, S., Maramba, I.D.C.,
Jones, K.H., & Ford, D.V. Sources of discovery, reasons for registration, and expectations of an
internet-based register for Multiple Sclerosis: Visualisations and explorations of word uses and
contexts. International Journal of Healthcare Information Systems and Informatics 2012; 7: 30-
46.

5.   Osborne, L.A., Lockhart-Jones, H.M., Middleton, R.M., Thompson, S., Maramba, I.D.C.,
Jones, K.H., Ford, D.V., & J.G. Noble (under review). Barriers to the use of an internet-register
for Multiple Sclerosis as identified by internet-using people with MS based on their experiences.

6.   http://www.themcfox.com/multiple-sclerosis/ms-facts/multiple-sclerosis-facts.htm;
http://www.atlasofms.org/export.aspx (accessed 10/03/2011).

7.   http://www.mssociety.org.uk/about_ms/index.html (accessed 10 Mar 2011.

8.   Alonso, A. & Hernan, M. Temporal trends in the incidence of multiple sclerosis: a
systematic review. Neurology 2008; 71: 129-135.

9.   Compston, A., & Coles, A. Multiple Sclerosis. Lancet 2008; 372: 1502-17.

10. Bol, Y., Duits, A.A., Lousberg, R. et al. Fatigue and physical disability in patients with multiple sclerosis: a structural equation modeling approach. Journal of Behavioral Medicine 2010; 33: 355–363.

11. O'Connor, A.B., Schwid, S.R., Herrmann, D.N., Markman, J.D., & Dworkin, R.H. Pain associated with multiple sclerosis: Systematic review and proposed classification. Pain 2008; 137: 96-111.

12. Sharrack, B., Hughes, R.A.C., Soudain, S., & Dunn, G. The psychometric properties of clinical rating scales used in multiple sclerosis. Brain 1999; 122: 141–159.

13. Amato, M.P., Zipoli, V., & Portaccio, E. Multiple sclerosis-related cognitive changes: A review of cross-sectional and longitudinal studies. Journal of the Neurological Sciences 2006; 245: 41-46.

14. Chiaravalloti, N.D., & DeLuca J. Cognitive impairment in multiple sclerosis. Lancet Neurology 2008; 12: 1139-51.

15. Patti, F., Amato, M.P., Trojano, M. et al. Cognitive impairment and its relation with disease measures in mildly disabled patients with relapsing-remitting multiple sclerosis: baseline results from the Cognitive Impairment in Multiple Sclerosis (COGIMUS) study. Multiple Sclerosis 2009; 15: 779-788.

16. Dennison, L., Moss-Morris, R., & Chalder, T. A review of psychological correlates of adjustment in patients with multiple sclerosis. Clinical Psychology Review 2009; 29: 141-153.

17. Devins, G.M., & Seland, T.P. Emotional impact of multiple sclerosis: recent findings and suggestions for future research. Psychological Bulletin 1987; 101: 363-375.

18. Kalb, R. The emotional and psychological impact of multiple sclerosis relapses. Journal of Neurological Science 2007; 256 Suppl 1: S29-33.

19.   Benito-Leon, J.N., Morales, J.M. Rivera-Navarro, J.S., & Mitchel, A.J. A review about the impact of multiple sclerosis on health-related quality of life Disability and Rehabilitation 2003; 25: 1291-1303.

20.   Mitchell, A.J., Benito-León, J., Morales González, J-M., & Rivera-Navarro, J. Quality of life and its assessment in multiple sclerosis: integrating physical and psychological components of wellbeing. Lancet Neurology 2005; 4: 556–66.

21.   Motl, R.W., McAuley, E., Snook, E.M., & Gliottoni, R.C. Physical activity and quality of life in multiple sclerosis: Intermediary roles of disability, fatigue, mood, pain, self-efficacy and social support. Psychology and Health Medicine 2009; 14: 111–124.

22.   Cohen JA, Ringold SC, Polman CH, Wolinksky JS. Disability outcome measures in multiple sclerosis clinical trials: current status and future prospects. Lancet Neurology 2012; 11: 467-76

23.   Noble, J.G., Osborne, L.A., Jones, K.H., Middleton, R.M., & Ford, D.V. Commentary on: "Disability Outcome Measures in Multiple Sclerosis Clinical Trials". *Multiple Sclerosis Journal* 2012; 18.

24.   Ebers, G.C. Commentary: Outcome measures were flawed. British Medical Journal 2010; 340: 2693.

25.   Cofield, S.S., & Cutter, G.R. Outcome measures in multiple sclerosis. In C.S. Raine, H.F. McFarland, & R. Hohfeld (Eds.), Multiple Sclerosis: a comprehensive text. Philadelphia: Elsevier.

26.   National Institute for Clinical Excellence Multiple Sclerosis. Management of multiple sclerosis in primary and secondary care. 2003**.** London. National Institute for Clinical Excellence.

27.   Amato, M.P., *&* Portaccio*,* E. Clinical outcome measures in multiple sclerosis. Journal of Neurological Science 2007; 259: 118-122.

28.   Bandari, D.S., Vollmer, T.L., Khatri, B.O., & Tyry, T. Assessing Quality of Life in Patients with Multiple Sclerosis. International Journal of MS Care 2010; 12: 34-41.

29.   Benedict, R.H.B., & Zivadinov, R. Reliability and validity of neuropsychological screening and assessment strategies in MS. Journal of Neurology 2007; 254 [Suppl 2]: II/22–II/25.

30.   D'Souza, M., Kappos, L., & Czaplinski A. Reconsidering clinical outcomes in Multiple Sclerosis: relapses, impairment, disability and beyond. Journal of the Neurological Sciences 2008; 274: 76-79.

31.   Kragt, J.J., Thompson, A.J., Montalban, X. et al. Responsiveness and predictive value of EDSS and MSFC in primary progressive MS. Neurology 2008; 70: 1084-1091.

32.   Scherer, P. Cognitive screening in multiple sclerosis. Journal of Neurology 2007; 254 [Suppl 2]: II/26–II/29.

33.   Stone, L.A. Getting the measure of multiple sclerosis. The Lancet Neurology; 8: 221-222.

34.   Whitehead, L. The measurement of fatigue in chronic illness: A systematic review of unidimensional and multidimensional fatigue measures. Journal of Pain and Symptom Management; 37: 107-128.

35.   Bethoux F, Bennett S. Evaluating walking in patients with Multiple Sclerosis: Which assessment tools are useful in a clinical practice. International Journal of MS Care 2011; 13: 4-14.

36.   Kragt JJ, Nielsen JM, van der Linden FAH, Polman CH, Uitdehagg BMJ. Disease progression in multiple sclerosis: combining physicians and patients perspectives? Multiple Sclerosis Journal 2011; 17: 234-40.

37.   Rosova K, Martinkova P, Vyskotova J, Sedova M. Assessment set for evaluation of clinical outcomes in multiple sclerosis: psychometric properties. Patient Related Outcome Measures 2012; 3: 59-70.

38.   Hyland M, Rudick RA. Challenges to clinical trials in multiple sclerosis: outcome measures in the era of disease modifying drugs. Current Opinion in Neurology 2011; 24: 255-61.

39.   Bohannon, R.W., & Smith, M.B. Interrater reliability of a Modified Ashworth Scale of muscle spasticity. Physical Therapy 1987; 67: 207-206.

40.   Ashworth, B. Preliminary trial of carisprodal in multiple sclerosis. Practitioner 1964; 192: 540-542.

41.   Paltamaa, J.W.H., Sarasoja, T., Wikstrom, J., & Malkia, E. Reliability of physical functioning measures in ambulatory subjects with MS. Physiotherapy Research International 2005; 10: 93-109.

42.   Nuyens G, De Weerdt W, Ketelaer P *et al*. Inter rater reliability of the Ashworth Scale in multiple sclerosis. Clin Rehabil 1994; **8**: 286–92.

43.   Ansari, N.N., Naghdi, S., Moammeri, H., & Jalaie, S. Ashworth Scales are unreliable for the assessment of muscle spasticity. Physiotherapy Theory and Practice 2006; 22: 119-125.

44.   Pandyan, A.D., Johnson, G.R., Price, C.I., Curless, R.H., Barnes, M.P., & Rodgers H. A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity. Clinical Rehabilitation 1999; 13: 373-83.

45.   Farrar, J.T., Troxel, A.B., Stott, C., Duncombe, P., & Jensen, M.P. Validity, reliability, and clinical importance of change in a 0-10 numeric rating scale measure of spasticity: a post hoc analysis of a randomized, double-blind, placebo-controlled trial. Clinical Therapy 2008; 30: 974-85.

46.   Cooper, A., Musa, I.M., & Wiles, C.M. Electromyography characterization of stretch responses in hemiparetic stroke patients and their relationship with the Modified Ashworth Scale. Clinical Rehabilitation 2005; 19: 760-766.

47.   Katz, R., Rovai, G.P., Brait, C., Rymer, W.Z. Objective quantification of spastic hypertonia: correlation with clinical findings. Archive of Physical Medicine Rehabilitation 1992; 73: 339-347.

48.   Fleuren, J.F.M., Voerman, G.E., Erren-Wolters, C.V. et al. Stop using the Ashworth Scale for the assessment of spasticity**.** Journal of Neurology, Neurosurgery, and Psychiatry 2010; 81**:** 46-52.

49.   Ghotbi N, Ansari NN, Naghdi S, Hasson S. Measurement of lower-limb muscle spasticity: Intrarater reliability of Modified Modified Ashworth Scale. Jouranl of Rehabilitation Research and Development 2011; 48: 83-8.

50.   Kurtzke, J.F. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983; 33: 1444–52.

51.   Ingram, G., Colley, E., Ben-Shlomo, Y., et al. Validity of patient-derived disability and clinical data in multiple sclerosis. Multiple Sclerosis 2010; 16: 472-479.

52.   Hadjimichael, O., Kerns, R.D., Rizzo, M.A., Cutter, G., & Vollmer, T. Persistent pain and uncomfortable sensations in persons with multiple sclerosis. Pain 2007; 127: 35-41.

53.   Amato, M.P., Fratiglioni, L., Groppi, C., Siracusa, G., & Amaducci, L. Interrater reliability in assessing functional systems and disability on the Kurtzke scale in multiple sclerosis. Archives of Neurology 1988; 45: 746–8.

54.   Goodkin, D.E., Cookfair, D., Wende, K., et al. Inter- and intrarater scoring agreement using grades 1.0 to 3.5 of the Kurtzke Expanded Disability Status Scale (EDSS). Neurology 1992; 42: 859–63.

55.   Hobart, J., Freeman, J., & Thompson, A. Kurtzke scales revisited: the application of psychometric methods to clinical intuition. Brain 2000; 123: 1027–40.

56.   Polman, C.H., & Rudick, R.A. The Multiple Sclerosis Functional Composite: A clinically meaningful measure of disability. Neurology 2010; 74: S8–15.

57.   Van Winsen, L.M.L., Kragt, J.J., Hoogervorst, E.L.J., Polman, C.H., Uitdehaag, B.M.J. Outcome measurement in multiple sclerosis: detection of clinically relevant improvement. Multiple Sclerosis 2010; 16: 604-610.

58.   Hobart, J.C., Lamping, D.L., Freeman, J.A., & Thompson, A.J. Reliability, validity and responsiveness of the Kurtzke expanded disability status scale (EDSS) in multiple sclerosis (MS) patients. European Journal of Neurology 1996; 3: Suppl 4: 13.

59.   Hohol, M.J., Orav, E.J., & Weiner, H.L. Disease steps in multiple sclerosis: a longitudinal study comparing disease steps and EDSS to evaluate disease progression. Multiple Sclerosis 1999; 5: 349-54.

60.   Ebers, G.C., Heigenhauser, L., MSc, Daumer, M. Lederer, C., & Noseworthy, J.H. Disability as an outcome in MS clinical trials. Neurology 2008; 71: 624-631.

61.   Zhang J, Waabant E, Cutter G, Wolinsky JS, Glanzman R. EDSS variability before randomization may limit treatment discovery in primary progressive MS. Multiple Sclerosis Journal 2012; 18.

62.   Goldman MD, Motl RW, Rudick RA. Possible clinical outcome measures for clinical trials in patients with multiple sclerosis. Therapeutic Advances in Neurological Disorders 2010; 3: 229–239.

63.   Keith, R.A., Granger, C.V., Hamilton, B.B., et al. The functional independence measure: a new tool for rehabilitation. In: Eisenberg M, Grzesiak R, eds. Advances in clinical rehabilitation. New York: Springer Verlag, 1987: 6-1

64.   Hinkle, J.L. McClaran, J. Davies, J. & Ng, D. Reliability and Validity of the Adult Alpha Functional Independence Measure Instrument in England Journal of Neuroscience Nursing 2010; 42: 12-18.

65.   Turner-Stokes, L., Nyein, K., Turner-Stokes, T., & Gatehouse, C. The UK FIM+FAM: development and evaluation. Clinical Rehabilitation 1999; 13: 277-87.

66.   Brosseau, L., & Wolfson, C. The interrater reliability and the concurrent validity of the Functional Independence Measure (FIM) for Multiple Sclerosis patients. Journal of Clinical Rehabilitation 1994; 8: 107-115.

67.   Ottenbacher, K.J., Hsu, Y., Granger, C.V., & Fielder, R.C. (1996) The reliability of the FIM: a quantitative review for multiple sclerosis subjects Clinical Rehabilitation 1996; 8: 107-15.

68.   Segal, M.E., Gillard, M., & Schall R. Telephone and in-person proxy agreement between stroke patients and caregivers for the functional independence measures. American Journal of Physical Medicine Rehabilitation 1996; 75: 208-212.

69.   van der Putten, J.J., Hobart, J.C., Freeman, J.A., & Thompson, A.J. Measuring change in disability after inpatient rehabilitation: Comparison of the responsiveness of the Barthel Index and the Functional Independence Measure. Neurology Neurosurgery and Psychiatry 1999; 66: 480-484.

70.   Kidd, D., Stewart, G., Baldry, J., et al. The Functional Independence Measure: A comparative validity and reliability study. Disability Rehabilitation 1995; 17: 10-14.

71.   Wright, J. (2000). The FIM(TM). The Center for Outcome Measurement in Brain Injury. http://www.tbims.org/combi/FIM (accessed 6/6/2011).

72.   Sharrack, B., & Hughes, R.A.C. The Guy's Neurological Disability Scale (GNDS): a new disability measure for multiple sclerosis. Multiple Sclerosis 1999; 5: 223-233.

73.   Wicks P, Vaughan TE, Massagli MP. The Multiple Sclerosis Rating Scale revised (MSRS-R): Development, refinement, and psychometric validation using an online community. Health and Quality of Life Outcomes 2012; 10: 70.

74.   Mokkink LB, Knol DL, Uitdehaag BMJ. Factor structure of Guy's Neurological Disability Scale in a sample of Dutch patients with multiple sclerosis. Multiple Sclerosis 2011; 17: 1498-1503.

75.   Rossier, P., & Wade, D.T. The Guy's Neurological Disability Scale in patients with multiple sclerosis: a clinical evaluation of its reliability and validity. Clinical Rehabilitation 2002; 16: 75–95.

76.   Nicholl, C.R., Lincoln, N.B., Francis, V.F., & Stephan, T.F. Assessment of emotional problems in people with multiple sclerosis. Clinical Rehabilitation 2001; 15: 657.

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

78.   Fisher, J.S., Jak, A.J., Kniker, J.E., Rudick, R.A., & Cutter, G. Multiple sclerosis functional composite (MSFC): Administration and scoring manual – Revised. National Multiple Sclerosis Society 2001.

79.   Rudick, R.A., Cutter, G. & Reingold, S. The multiple sclerosis functional composite: a new

clinical outcome measure for multiple sclerosis trials. Multiple Sclerosis 2002*; 8: 359–365.*

80.   Solari, A., Radice, D., Manneschi, L., Motti, L., & Montanari, E. The multiple sclerosis

functional composite: different practice effects in the three test components. Journal of

Neurological Science 2005; 228*: 71–74.*

81.   Cohen, J.A., Fischer, J.S., Bolibrush, D.M., et al. Intrarater and interrater reliability of the

MS functional composite outcome measure. Neurology 2000; 54: 802-806.

82.   Cohen, J.A., Cutter, G.R., Fischer, J.S., et al. Use of the Multiple Sclerosis Functional

Composite as an Outcome Measure in a Phase 3 Clinical Trial. Archives of Neurology 2001; 58:

961-967.

83.   Miller, D.M., Rudick, R.A., Cutter, G., Baier, M., & Fischer, J. Clinical significance of the

Multiple Sclerosis Functional composite: Relationship to patient-reported Quality of life.

Archives of Neurology 2000; 57: 1319-1324.

84.   Portaccio E, Goretti B, Zipoli V, et al. Reliability, practical effects, and change indices for

Roa's brief repeatable battery. Multiple SCletosis 2010; 16: 611-17.

85.   Kalkers, N.F., De Groot, V., Lazeron, R.H., et al. Multiple sclerosis functional composite:

relation to disease phenotype and disability strata. Neurology 2000; 54: 1233–9.

86.   Hobart, J.C., Riazi, A., Lamping, D.L., Fitzpatrick, R., & Thompson, A.J. Improving the

evaluation of therapeutic interventions in multiple sclerosis: development of a patient-based

measure of outcome. Health Technologies and Assessment 2004; 8: 1–60.

87.   de Groot, V., Beckerman, H., Uitdehaag, B.M.J., et al. The usefulness of evaluative outcome

measures in patients with multiple sclerosis. Brain 2006; 129: 2648-2659.

88.   Bosma, L.V., Kragt, J.J., Brieva, L., et al. Progression on the Multiple Sclerosis Functional

Composite in multiple sclerosis: what is the optimal cut-off for the three components? Multiple

Sclerosis 2010; 16 :862-7.

89.   Graham, R.C., & Hughes, R.A.C. Clinimetric properties of a walking scale in peripheral

neuropathy. Journal of Neurology Neurosurgery, and Psychiatry 2006; 77: 977–979.

90.   Motl, R.W., & Snook, E.M. Confirmation and extension of the validity of the Multiple

Sclerosis Walking Scale-12 (MSWS-12). Journal of the Neurological Sciences 2008; 268: 69-73.

91.   Motl RW, McAuley E, Mullen S. Longitudinal measurement invariante of the Multiple

Sclerosis Walking Scale-12. Journal of Neurological Sciences 2011; 305: 75-9.

92.   Spain RL, George RJ, Salarian A, et al. Body-worn motion sensors detect balance and gait

deficits in people with multiple sclerosis who have normal walking speed. Gait & Posture 2012;

4: 573-8.

93.   Snook, E.M., Motl, R.W., & Gliottoni, R.C. The effect of walking mobility on the

measurement of physical activity using accelerometry in multiple sclerosis. Clinical

Rehabilitation 2009; 23: 248.

94.   Pilutti LA, Dlugonski D, Sandroff BM, et al. Further validation of Multiple Sclerosis

Walking Scale-12 score based on spatiotemporal gait parameters. Archives of Physical Medicine

and Rehabilitation 2012.

95.   Holland, A., O'Connor, R.J., Thompson, A.J., Playford, E.D., & Hobart, J.C. Talking the

talk on walking the walk: a 12-item generic walking scale suitable for neurological conditions?

Journal of Neurology 2006; 253: 1594-602.

96.   Motl RW, Mullen S, McAuley E. Multigroup measurement invariance of the multiple

sclerosis wlaking scale-12. Neurological Research 2012; 34: 149-52.

97.   Folstein, M.F., Folstein, S.E., & McHugh, P.R. "Mini-mental state". A practical method for

        grading the cognitive state of patients for the clinician. Journal of psychiatric Research 1975; 12:

        189-198.

98.   Molloy, D.W., Alemayehu, E., & Roberts, R. Reliability of a Standardized Mini-Mental

        State Examination compared with the traditional Mini-Mental State Examination. American

        Journal of Psychiatry 1991; 148: 102-5.

99.   Beatty, W.W. & Goodkin, D.E. Screening for cognitive impairment in multiple sclerosis. An

        evaluation of the Mini-Mental State Examination. Archives of Neurology 1990; 47: 297–301.

100.   Swirsky-Sacchetti, T., Field, H.L., Mitchell, D.R., et al. The sensitivity of the Mini-Mental

        State Exam in the white matter dementia of multiple sclerosis. Journal of Clinical Psychology

        1992; 48: 779-86.

101.   Scherer, P. Cognitive screening in multiple sclerosis. Journal of Neurology 2007; 254 [Suppl

        2]: II/26–II/29.

102.   Bowie, P., Branton, T., & Holmes, J. Should mini mental state examination be used to

        monitor dementia treatments? Lancet 1999; 354: 1527–1528

103.   Rogers, J.M., & Panegyres, P.K. Cognitive impairment in multiple sclerosis: Evidence-based

        analysis and recommendations. Journal of Clinical Neuroscience 2007; 14: 919-927.

104.   Hosseini H, Mandat T, Waabaat E, et al. Unilateral thalamic deep brain stimulation for

        disabling kinetic tremor in multiple sclerosis. Neurosurgery 2012; 70: 66-69.

105.   Beck, A.T., Brown, G.K., & Steer, R.A. Beck Depression Inventory-II (BDI-II). 1996. San

        Antonio, TX. The Psychological Corporation.

106.   Beck, A.T., Ward, C.H., Mendelson, M., et al. An inventory for measuring depression.

        Archives of General Psychiatry 1961; 4: 53-63.

107.   Aikens, J.E., Reinecke, M.A., Pliskin, N.H., et al. Assessing Depressive Symptoms in

Multiple Sclerosis: Is It Necessary to Omit Items from the Original Beck Depression Inventory?

Journal of Behavioral Medicine 1999; 22: 127-142

108.   **A**rnau, R.C.**,** Meagher, M.W., Norris, M., & Bramson, R. Psychometric evaluation of the

Beck Depression Inventory-II with primary care medical patients. Health Psychology 2001; 20:

112-119.

109.   Moran, P.J., & Mohr, D.C. The Validity of Beck Depression Inventory and Hamilton Rating

Scale for Depression Items in the Assessment of Depression Among Patients with Multiple

Sclerosis. Journal of Behavioral Medicine 2005; 28: 35-41.

110.   Viinamaki, H., Tanskanen, A., Honkalampi, K., et al. Is the Beck Depression Inventory

suitable for screening major depression in different phases of the disease? Nordic Journal of

Psychiatry 2004; 58: 49-53.

111.   Lewis-Rosenbaum, D. Factors associated with severity of depression in multiple sclerosis.

Ph.D. thesis, 2005, Yeshiva University.

112.   Benedict, R.H., Fishman, I., McClellan, M.M., Bakshi, R., & Weinstock-Guttman, B.

Validity of the Beck Depression Inventory-Fast Screen in multiple sclerosis. Multiple Sclerosis

2003; 9: 393-6.

113.   Siegert, R.J., Tennant, A., & Turner-Stokes, L. Rasch analysis of the Beck Depression

Inventory-II in a neurological rehabilitation sample. Disability and Rehabilitation 2010; 32: 8-17.

114.   Goldberg, D.P. & Williams, P. The User's Guide to the General Health Questionnaire. 1988.

Windsor: NFER—Nelson.

115.   Hobart, J.C., Riazi, A., Thompson, A.J., et al. Getting the measure of spasticity in multiple

sclerosis: the Multiple Sclerosis Spasticity Scale (MSSS-88). Brain 2006; 129**:** 224-234.

116. Vallejo, M.A., Jordán, C.M., Díaz, M.I., Comeche, M.I., & Ortega, J. Psychological Assessment via the Internet: A Reliability and Validity Study of Online (vs Paper-and-Pencil) Versions of the General Health Questionnaire-28 (GHQ-28) and the Symptoms Check-List-90-Revised (SCL-90-R). Journal of Medical Internet Research 2007; 9:e2.

117. Ormel, J., Koeter, M.W.J., van den Brink, W., & Giel, R. Concurrent validity of GHQ-28 and PSE as measures of change. Psychological Medicine 1989; 19: 1007-1013.

118. Robins. P.V., & Brooks. B.R. Emotional disturbance in multiple sclerosis patients: Validity of "GHQ". Psychological Medicine 1981; 11: 425-427.

119. Goldberg, D.P. Identifying psychiatric illness among general medical patients. British Medical Journal 1985; 291: 161–163.

120. Banks, M. Validation of the General Health Questionnaire in a young community sample. Psychological Medicine 1983; 13: 349-353.

121. Meader N, Mitchell AJ, Chew-Graham C, et al. Case identification of depression in patients with chronic physical health problems: a diagnostic accuracy meta-analysis of 113 studies. British Journal of General Practice 2011; 61: 808-20.

122. Hobart, J.C., Riazi, A., Lampling, D.L., Fitzpatrick, R. & Thompson, A.J. How responsive is the Multiple Sclerosis Impact Scale (MSIS-29)? A comparison with some other self report scales. Journal of Neurology Neurosurgery and Psychiatry 2005; **76:** 1539-1543.

123. Freeman, J.A., Langdon, D.W., Hobart, J.C., et al. Clinical appropriateness: a key factor in outcome measure selection: the 36 item short form health survey in multiple sclerosis. J Neurol Neurosurg Psychiatry 2000; 68: 150–156.

124. Zigmond, A.S., & Smith R.P. The hospital anxiety and depression scale. Acta Psychiatrica Scandinavica 1983; 67: 361-370.

125.   Jones KH, Ford DV, Jones PA, John A, Middleton RM, Lockhart-Jones H, Osborne LA,

Noble JG. A large-scale study of anxiety and depression in people with Multiple Sclerosis: A

survey via the web portal of the UK MS Register 2012; *PlosONE*: 7.

126.   Janssens, A.C.J.W., van Doorn, P.A., de Boer, J.B., et al. Anxiety and depression influence

the relation between disability status and quality of life in multiple sclerosis. Multiple Sclerosis

2003; 9: 397-403.

127.   Spinhoven, P., Ormel, J., Sloekers, P.P., Kempen, G.I., Speckens, A.E., & Van Hemerts,

A.M. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different

groups of Dutch subjects. Psychological Medicine 1997; 27: 363–70.

128.   Bjelland, I., Dahl, A.A., Haug, T.T., & Neckelman, D. The validity of the Hospital Anxiety

and Depression Scale: An updated literature review. Journal of Psychosomatic Research 2002;

52: 69-77.

129.   Honarmand, K., & Feinstein, A. Validation of the Hospital Anxiety and Depression Scale

for use with multiple sclerosis patients. Multiple Sclerosis 2009; 15: 1518-1524.

130.   Gusi, N., Olivares, P.R., & Rajendram, R. The EQ-5D quality of life questionnaire. In

Preedy, V (ed) Handbook of Disease Burdens and Quality of Life Measures. Springer, London.

88-99, 2010.

131.   Dorman, P.J., Slattery, J., Farrell, B., Dennis, M. & Sandercock, P. Qualitative comparison

of the reliability of health status assessments with the EuroQol and SF-36 questionnaires after

stroke. Stroke 1998; 29: 63-68.

132.   Fisk, J., Brown, M., Sketris, I., Metz, L., Murray, T., & Stadnyk, K. A comparison of health

utility measures for the evaluation of multiple sclerosis treatments. Journal of Neurology

Neurosurgery and Psychiatry 2005; 76: 58-63.

133.   Ziemssen T. Symptom management in patients with Multiple Sclerosis. Journal of

Neurological Sciernces 2011; 311: S48-S52.

134.   McPherson, K., Myers, J., Taylor, W.J., McNaughton, H.K., & Weatherall, M. Self-

valuation and societal valuations of health state differ with disease severity in chronic and

disabling conditions. Medical Care 2004; 42: 1143-51.

135.   Hemmet, L., Holmes, J., Barnes, M., & Russell, N. What drives quality of life in multiple

sclerosis? Quarterly Journal of Medicine 2004; 97: 671-676.

136.   Kobelt, G., Berg, J., Lindgren, P., Fredrikson, S., & Jönsson, B. Costs and quality of life of

patients with multiple sclerosis in Europe. Neurologia I Neurochirurgia Polska 2006; 77: 58-365.

137.   Versteegh MM, Leunis A, Uyl-de-Groot CA, Stalk E. Condition-specific preference-based

measures: benefit or burden. Value in Health 2012; 15: 504-13.

138.   Fisk, J.D., Ritvo, P.G., Ross, L., Haase, D.A., Marrie, T.J., & Schlech, W.F. Measuring the

functional impact of fatigue: Initial validation of the fatigue impact scale. Clinical Infectious

Diseases 1994; 18(Suppl. 1): S79–S83.

139.   Fisk, J.D., Pontefract, A., Ritvo, P.G., Archibald, C.J., & Murray, T.J. The impact of fatigue

on patients with multiple sclerosis. Canadian Journal of Neurological Science 1994; 21: 9-14.

140.   Brown, R.F., & Schutte, N.S. Direct and indirect relationships between emotional

intelligence and subjective fatigue in university students. Journal of Psychosomatic Research

2006; 60: 585-593.

141.   Kos, D., Kerckhofs, E., Carrea, I., Verzo, R., Ramos, M., & Jansa, J. Evaluation of the

Modified Fatigue Impact Scale in four different European countries. Multiple Sclerosis 2005; 11:

76-80.

142.   Mathiowetz V. Test-retest reliability and convergent validity of the Fatigue Impact Scale for persons with multiple sclerosis. American Journal of Occupational Therapy 2003; 57: 389-95.

143.   Pittion-Vouyovitch, S., Debouverie, M., Guillemin, F., Vandenberghe, N, Anxionnat, R., & Vespignani, H. Fatigue in multiple sclerosis is related to disability, depression and quality of life. Journal of the Neurological Sciences 2006; 243: 39-45.

144.   Rietberg, M.B., Van Wegen, E.E.H., & 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: 1870-1876.

145.   Tellez, N., Rıo, J., Tintore, M., Nos, C., Galan, I., & Montalban, X. Does the Modified Fatigue Impact Scale offer a more comprehensive assessment of fatigue in MS? Multiple Sclerosis 2005; 11: 198-202.

146.   Flachenecker, P., Kümpfel, T., Kallmann, B., et al. Fatigue in multiple sclerosis: a comparison of different rating scales and correlation to clinical parameters. Multiple Sclerosis 2002; 8: 523-526.

147.   Vanage, S.M., Gilbertson, K.K., & Mathiowetz, V. Effects of an Energy Conservation Course on Fatigue Impact for Persons With Progressive Multiple Sclerosis. Occupational Therapy 2003; 57: 315-323.

148.   Bruce, J.M., Hancock, L.M., Arnett, P., & Lynch, S. Treatment adherence in multiple sclerosis: association with emotional status, personality, and cognition. Journal of Behavioral Medicine 2010; 33: 219–227.

149.   Herbert JR, Corby JR, Manajo MM, Schenkman M. Effect of vestibular rehabilitation on multiple sclerosis-related fatigue and upright postural control: A randomized control study. Physical therapy 2011; 91: 1166-83.

150. Mills, R.J., Young, C.A., Pallant, J.F., & Tennant, A. Rasch analysis of the Modified Fatigue Impact Scale (MFIS) in multiple sclerosis. Journal of Neurology Neurosurgery and Psychiatry 2010; 81: 1049-1051.

151. Doward, L.C., Meads, D.M., Fisk, J., et al. International development of the Unidimensional Fatigue Impact Scale (U-FIS). Value in Health 2010; 13: 463-8.

152. Multiple Sclerosis Council for Clinical Practice Guidelines (MSCCPG). (1998). Fatigue and multiple sclerosis: Evidence based management strategies for fatigue in multiple sclerosis. Washington, DC: Paralyzed Veterans of America.

153. Ford, H.L., Gerry, E., Tennant, A., Whalley, D., Haigh, R., & Johnson, M.H. Developing a disease specific quality of life measure for people with multiple sclerosis. Clinical Rehabilitation 2001; 15: 247-258.

154. Greenhalgh, J., Ford, H., Long, A., & Hurst, K. The MS Symptom and Impact Diary (MSSID): psychometric evaluation of a new instrument to measure the day to day impact of multiple sclerosis. Journal of Neurology Neurosurgery and Psychiatry 2004; 75: 577-582.

155. Learmonth YC, Paul L, Miller L, Mattison P, McFadyen AK. The effect of a 12 week leisure centre-based, group exercise intervention for people moderately effected with multiple sclerosis: a randomized controlled pilot study. Clinical Rehabilitation 2012; 26: 579-93.

156. Jongen, P.J. Lehnick, D., Sanders, E., et al. Health-related quality of life in relapsing remitting multiple sclerosis patients during treatment with glatiramer acetate: a prospective, observational, international, multi-centre study. Health and Quality of Life Outcomes 2010; 8: 133.

157. Vickrey, B.G., Hays, R.D., Harooni, R., Myers, L.W., & Ellison, G.W. A health related quality of life measure for multiple sclerosis. Quality of Life Research 1995; 4: 187–206.

158.   Hobart, J., Lamping, D., Fitzpatrick, R., Riazi, A., & Thompson, A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. Brain 2001; 124: 962-73.

159.   Ramp, M., Khan, F., Misajon, R.A., & Pallant, J.F. Rasch analysis of the Multiple Sclerosis Impact Scale MSIS-29. Health Quality of Life Outcomes 2009; 7: 58.

160.   Riazi, A., Hobart, J., Lamping, D., Fitzpatrick, R., & Thompson, A. Multiple Sclerosis Impact Scale (MSIS-29): reliability and validity in hospital based samples. Journal of Neurology Neurosurgery and Psychiatry 2002; 73: 701–704.

161.   Smedal, T., Johansen, H.H., Myhr, K-M., & Strand, L.I. Psychometric properties of a Norwegian version of Multiple Sclerosis Impact Scale (MSIS-29). Acta Neurol Scand 2010; 122: 244–251.

162.   van der Linden, F.A.H., Kragt, J.J., Klein, M., van der Ploeg, H.M., Polman, C.H., & Uitdehaag, B.M.J. Psychometric evaluation of the multiple sclerosis impact scale (MSIS-29) for proxy use. Journal of Neurology Neurosurgery and Psychiatry 2005; 76: 1677-1681.

163.   Motl RW, McAuley E, Wynn D, Sandroff B, Soh Y. Physical activity, self-efficacy and health-related quality of life in people with multiple sclerosis: analysis of associations between individual-led change over a year. Quality of Life Research 2012.

164.   Hoogervorst, E.L.J, Zwemmer, J.N.P., Jelles, B., Polman, C.H., & Uitdehaag, B.M.J. Multiple Sclerosis Impact Scale (MSIS-29): relation to established measures of impairment and disability. Multiple Sclerosis 2004; 10: 569-74.

165.   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: 841–844.

166.   Giordano, A., Pucci, E., Naldi, P., et al. Responsiveness of patient reported outcome

measures in multiple sclerosis relapses: the REMS study. Journal Neurology Neurosurgery and

Psychiatry 2009; 80: 1023-8.

167.   Heiskanen, S., Meriläinen, P., & Pietilä, A-M. Health-related quality of life - testing the

reliability of the MSQOL-54 instrument among MS patients. Scandinavian Journal of Caring

Sciences 2007; 21: 199-206.

168.   Ghaem, H., Borhani Haghighi, A., Jafari, P., & Nikseresht, A,R. Validity and reliability of

the Persian version of the multiple sclerosis quality of life questionnaire. Neurology India 2007;

55: 369-375.

169.   Yamamoto, T., Ogata, K., Katagishi, M., et al. Validation of the Japanese-translated version

Multiple Sclerosis Quality of Life-54 instrument. Rinsho Shinkeigaku 2004; 44: 417-21.

170.   Nicholl, L., Hobart, J.C., Cramp, A.F.L., & Lowe-Strong, A.S. Measuring quality of life in

multiple sclerosis: not as simple as it sounds. Multiple Sclerosis 2005; 11: 708-712.

171.   Drulovic J, Bursac LO, Milejkovic D, et al. MSQoL-54 predicts change in fatigue alter

inpatient rehabilitation for people with múltiple sclerosis. Disability and Rehabilitation 2012.

172.   Corallo F, De Luca, Leonardi R, De Salvo S, Bramanti P, Marino S. The clinical use of

quality of life scales in neurological disorders. Acta Neuropsychiatrica 2012.

173.   Stewart, A.L., Hays, R.D., & Ware, J.E. The MOS Short-form General Health Survey:

reliability and validity in a patient population. Medical Care 1988; 26: 724-35.

174.   Vickrey, B.G., Hays, R.D., Genovese, B.J., Myers, L.W., & Ellison, G.W. Comparison of a

generic to disease-targeted health-related quality-of-life measures for multiple sclerosis. Journal

of Clinical Epidemiology 1997; 50: 557-69.

175. Wynia, K., Middel, B., van Dijk, J.P., De Keyser, J.H., & Reijneveld, SA. The impact of disabilities on quality of life in people with multiple sclerosis. Multiple Sclerosis 2008; 14: 972-80.

176. Pfennings, L.E.M.A., van der Ploeg, H.M., Cohen, L., et al. A health-related quality of life questionnaire for multiple sclerosis Patients. Acta Neurologica Scandinavica; 100: 148-155.

Figure 1: Psychometric characteristics

Property	Measured	Assesses
Internal Reliability	Chronbach's Alpha	The degree to which all items are related to one another
Test-Retest Reliability	Correlation between test administered at two separate times	The freedom from unsystematic error over time
Concurrent Validity	Correlation between measure and other measures in the area	Whether the scale compares well with other similar tools
Sensitivity	Test-identified positives divided by all actual positives	The degree that the measure identifies individuals with the disorder (true positives)
Specificity	Test-identified negatives divided by all actual negatives	The degree to which the measure does not identify people without the disorder (true negatives)
Responsiveness	Effect sizes, or area under the receiver operating characteristics curve	The ability to detect changes in the measure over time

Table 2: Characteristics of measures.

Measure	Generic or MS Specific	Clinician Administered or Patient Completed	Approximate Length of Time to Administer or Complete	Domains Covered
Modified Ashworth Scale	Generic	Clinician	Short (depending on number of joints assessed)	Physical Disability: Muscle spasticity, joint by joint
Beck Depression Inventory	Generic	Patient	10 min	Psychological/Emotional: Symptoms of depression (e.g., hopelessness, irritability, guilt, and somatic problems)
Euro Qol - 5 Dimension Questionnaire	Generic	Patient	5 min	Quality of Life: Mobility, self-care, usual activities, pain, anxiety/depression, and general health.
Expanded Disability Status Scale	Generic (mainly used for MS)	Clinician administered and rated	Variable (depending on neurological examination)	Physical Disability: Impairment and disability (with 8 additional scales: pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual, cerebral, and other functioning)
Fatigue Impact Scale and Modified	Generic (developed on an MS population)	Patient	20 min	Quality of Life: Overall impact of fatigue, and sub-scales relating to the: physical, cognitive, and psychosocial, impacts of fatigue
Functional Independence Measure	Generic	Clinician or trained layperson	30 min	Quality of Life: Independence and ability of a patient to conduct routine daily activities concerned with motor function (e.g., self-care, sphincter control) and cognitive function (primarily social and communication function)

General Health Questionnaire	Generic	Patient	Up to 8 min (depending on version used)	Psychological/Emotional: General psychiatric problems (GHQ-28 scores somatic problems, anxiety, social dysfunction, and severe depression)
Guy's Neurological Disability Scale	MS Specific	Clinician or Patient	30 min	Physical Disability: Overall disability (12 sub-scales: cognition, mood, vision, speech, swallowing, upper and lower limb function, bladder and bowel function, sexual function, fatigue, and other)
Hospital Anxiety and Depression Scales	Generic	Patient	5 to 10 min	Psychological/Emotional: Anxiety and Depression
Leeds Multiple Sclerosis QoL Scale	MS Specific	Patient	5 min	Quality of Life: Overall index of quality of life, mostly focuses on fatigue and social issues
Mini Mental State Examination	Generic	Clinician	20 min	Cognitive Impairment: Cognitive impairment across a number of domains (e.g., attention, memory, orientation, arithmetic)
Multiple Sclerosis Functional Composite	MS Specific	Clinician	Variable (depending on disability level)	Physical/Cognitive Impairment: Impairment across 3 domains (leg function, arm function, and cognitive function)
Multiple Sclerosis Impact Scale	MS Specific	Patient	10 min	Quality of Life: Impact of MS on daily living on both physical and psychological aspects
Multiple Sclerosis QoL-54	MS Specific (generic component)	Patient	11 to 18 min according to Vickrey et al. 1995 (at least 30 min according to Bandari et	Quality of Life: Overall quality of life with two domains Physical Health (physical function, health perceptions, energy/fatigue, role limitations - physical, pain, sexual function, social function, and health distress); and Mental

				al., 2010)	Health (health distress, overall quality of life, emotional well-being, role limitations - emotional, and cognitive function)
Multiple Sclerosis Walking Scale	MS Specific (has been used for other disorders)	Patient	5 min	Physical Impairment: Disruption to walking and walking quality
Health Scale Short Form-36	Generic	Patient	10 to 15 min	Quality of Life: Overall impact of a disorder on functioning (with sub-scales for physical health and mental health, which, themselves, are based on a number of sub-scales that measure: physical functioning, role limitation, pain, general health, vitality, social functioning, and mental health.

Table 3: Quality of the psychometric properties of measures.

Measure	Internal Reliability	Test-Retest Reliability	Concurrent Validity	Sensitivity	Specificity	Responsive-ness
Modified Ashworth Scale	Grey	Green	Orange	White	White	White
Beck Depression Inventory	Green	Green	Orange	Green	Orange	Orange
Euro Qol - 5 Dimension Questionnaire	Grey	Orange	Orange	Red	White	Red
Expanded Disability Status Scale	Green	Orange	Green	Red	Green	Red
Modified/ Fatigue Impact Scale	Green	Green	Orange	Green	Green	Orange
Functional Independence Measure	Green	Green	Green	White	White	Orange
General Health Questionnaire	Green	Orange	Orange	Orange	Green	Orange
Guy's Neurological Disability Scale	Green	Green	Orange	White	White	Orange
Hospital Anxiety and Depression Scales	Green	Green	Orange	Orange	Green	Green

	1	2	3	4	5	6
Leeds MS QoL Scale	Green	Green	Yellow	White	White	Yellow
Mini Mental State Examination	Green	Yellow	Yellow	Red	Green	Red
MS Functional Composite	Green	Green	Green	Yellow	Green	Yellow
Multiple Sclerosis Impact Scale	Green	Green	Yellow	Green	Yellow	Yellow
Multiple Sclerosis QoL-54	Green	Yellow	Yellow	White	White	Yellow
Multiple Sclerosis Walking Scale	Green	Green	Green	Yellow	White	Green
Health Scale Short Form-36	Green	Yellow	Yellow	White	White	Red

Key:	
(Gray)	Not Applicable
(Red)	Poor (< 0.4)
(Yellow)	Moderate (0.4-0.7) /Mixed or Varied Values/Some Caution Required

	Good (> 0.7)
	No Data Available