CONSENSUS STATEMENT

Toward a Consensus on Rehabilitation Outcomes in MS: Gait and Fatigue

Report of a CMSC Consensus Conference, November 28–29, 2007

Brian Hutchinson, PT, MSCS; Susan J. Forwell, PhD, (OT)C, FCAOT; Susan Bennett, PT, EdD, NCS, MSCS; Theodore Brown, MD, MPH; Herb Karpatkin, PT, MSCS; Deborah Miller, PhD, LISW

A multidisciplinary consensus conference was held on November 28–29, 2007, by the Consortium of Multiple Sclerosis Centers (CMSC) to determine the most appropriate outcome measures for gait and fatigue in people with multiple sclerosis (MS). The goals of this conference were to 1) improve understanding of gait and fatigue outcome measures being used by rehabilitation professionals treating people with MS; 2) establish consensus on outcome measures; and 3) establish consensus on required follow-up for transfer of this knowledge to rehabilitation professionals. The consensus conference and this document are the initial steps toward achieving the stated goals. Although many measures of fatigue exist, it was recommended that a global outcome measure for fatigue be developed that would 1) include a screen for the functional ramifications of fatigue for activities and participation; 2) be quick and easy to administer; 3) demonstrate psychometric integrity for MS; and 4) examine fatigue over a continuum of the MS disease course. In addition, it was recommended that an assessment battery for fatigue be developed. With respect to gait outcome measures, it was agreed that the following tools should be included in a preliminary chart for use in a clinical setting: the Timed 25-Foot Walk, Timed Up and Go test, Dynamic Gait Index, 6-Minute Walk, and self-reported 12-item Multiple Sclerosis Walking Scale. The global outcome fatigue measure and assessment battery are currently being developed, and work on a detailed gait outcome measures chart and additional research on commonly used gait outcome measures are in progress. ***Int J MS Care.*** **2009;11:67–78.**

Reaching a multidisciplinary consensus among rehabilitation professionals on functional outcome measures in multiple sclerosis (MS) is an enormous challenge. This task has evolved over time as a result of the concerted effort of many members of the Consortium of Multiple Sclerosis Centers (CMSC). The journey toward consensus on gait and fatigue outcomes began in 2004, when a group of rehabilitation professionals from the United States and Canada met in Chicago to discuss current rehabilitation practices for MS patients and how to best standardize functional outcome measures. A consensus statement recommending the implementation of standardized functional outcome tools for use in the rehabilitation of MS patients was created and adopted by the CMSC Board of Governors.[1] At that time, it was proposed that one or two functional outcome tools be added to the array of assessment tools already in use by rehabilitation professionals—such as the Fatigue Impact Scale (FIS) or the Modified Fatigue Impact Scale (MFIS) for fatigue and the Timed 25-Foot Walk

From The Heuga Center for Multiple Sclerosis, Edwards, CO, USA (BH); Department of Occupational Science and Occupational Therapy, University of British Columbia, Vancouver, British Columbia, Canada (SJF); Department of Neurology and Rehabilitation Medicine, University of Buffalo, Buffalo, NY, USA (SB); MS Center at Evergreen, Kirkland, WA, USA (TB); New York Neurorehabilitation Group, New York, NY, USA (HK); and Mellen Center for Multiple Sclerosis Research and Treatment, Cleveland Clinic, Cleveland, OH, USA (DM). Mr. Hutchinson is now with Acorda Therapeutics, Hawthorne, NY, USA. *Correspondence:* Brian Hutchinson, PT, MSCS, c/o The Heuga Center for Multiple Sclerosis, 27 Main Street, Suite 303, Edwards, CO 81632.

(T25FW) or the Timed Up and Go (TUG) test for gait (Table 1).

Using those recommendations as a basis, and in view of the broad range of issues managed by rehabilitation professionals, the CMSC focused discussions more specifically on scales examining fatigue and gait. The rationale for including both issues was the intuitive sense of their dynamic relationship. One of the conundrums in MS is that walking can cause fatigue, and fatigue affects walking ability. It is important to understand the various characteristics that contribute to the link between fatigue and gait anomalies, such as strength, sensory changes, ataxia, spasticity, energy expenditure, multitasking, balance, and the environment.

When fatigue guidelines were developed in 1998,[2] fatigue was divided into two broad categories: nonprimary fatigue and primary fatigue. Nonprimary fatigue was further broken down into two subtypes: non-MS fatigue and fatigue secondary to MS. Non-MS fatigue can be acute or chronic and encompasses situations that are separate from the disease, such as infection, emotional stress, comorbid medical conditions, medication use, depression, and sleep disorders. Fatigue secondary to MS was described as primarily resulting from mobility and/or respiratory impairments. These guidelines make an explicit association between mobility and fatigue, implying that the increased energy costs of impaired mobility must be addressed when managing fatigue.

Table 1. Recommendations of the 2004 consensus statement for functional outcome measures in MS

Rehabilitation discipline	Recommended functional outcome tools
Occupational therapy	Nine-Hole Peg Test Modified Fatigue Impact Scale OR Fatigue Impact Scale
Physical therapy	Berg Balance Test 25-Foot Timed Walk OR Timed Up and Go test
Speech-language pathology	Clinical Swallowing Assessment Boston Naming Test AND/OR Frenchay Dysarthria Test

Abbreviation: MS, multiple sclerosis.
Note: Presented at the Rehabilitation Research Interest Group Meeting, 2006 CMSC Annual Meeting, as "Standardized functional outcome tools should be used in rehabilitation of individuals with MS."

Primary MS fatigue, as described in the guidelines, is a diagnosis of exclusion, once "non-MS" and "secondary to MS" fatigue factors are managed. The guidelines defined primary MS fatigue as "significant fatigue that persists despite adjustment of medications and management of mobility issues as well as confounding medical problems such as depression and sleep disruptions."[2] It is likely that primary MS fatigue coexists with other factors affecting fatigue.

A 2008 study of 50 people with MS reported that 72% of respondents experienced fatigue that might be due to nonprimary MS fatigue sources.[3] Interestingly, 52% were found to have unresolved gait problems, and 20% were deconditioned. The authors noted that primary MS fatigue probably is still present, though masked by these and other factors.

Developing outcome measures to better identify the problems associated with fatigue and mobility issues in MS requires an understanding of the measures that are currently in use. Therefore, a survey of CMSC members was undertaken on this subject.

Survey of Outcome Measures

During the summer of 2007, a questionnaire was sent to members of the CMSC (n = 1700) with experience in MS rehabilitation to obtain rankings of commonly used functional outcome measures. The response (n = 36) was low; this may have been due to difficulties with the database. The questionnaires were sent to multiple people at a single facility, but they were often completed and returned by the facility with input from several rehabilitation team members. In some cases, five to seven therapists may have contributed to a single questionnaire, although this was considered only one response. In addition, for some clinics a specific individual could not be identified, so the mailing was directed to a "clinic director" or "clinic coordinator"; these questionnaires may have been misplaced. Finally, the individual identified for the mailing may not have had sufficient experience with functional outcome measures to complete the survey. Regardless, the responses obtained represented a good geographic cross-section as well as professional diversity: there were 21 facilities in Canada and the United States, which included 11 physical therapists, 6 occupational therapists, 3 nurses and nurse practitioners, 3 physiatrists, and 1 physician's assistant. The patient base averaged 884 MS patients annually, providing us with a broad

sampling of individuals and centers with experience in treating patients with MS. The survey gathered information on gait and fatigue measures currently in use, factors important in assessing gait and fatigue, and treatment strategies.

The survey results indicated that many different gait scales are currently in use, including the T25FW, observation, the Dynamic Gait Index (DGI), the TUG test, and the 12-item Multiple Sclerosis Walking Scale (MSWS-12). In addition, professionals rated the usefulness of these scales quite differently (Table 2).

The survey was also designed to solicit feedback regarding important factors to consider when assessing gait. Spasticity, strength, and fatigue were most often identified. Balance was seen as important, as well as equipment used, such as ankle-foot orthoses. Respondents also identified ambulatory aids, safety, endurance, and the environment as significant factors. Many participants commented on assessing the consistency of the gait pattern, the pattern itself, and the quality of that pattern.

Fatigue scales were also specifically addressed in the survey. As with gait, many different fatigue scales are in use, with the Fatigue Severity Scale (FSS) and FIS being the most common. Other scales mentioned were the MFIS, Brief Fatigue Index, and self-report. Although not a scale per se, interviews were also acknowledged by several individuals as important. As with gait, ratings for these scales varied widely (Table 2). Most participants identified the scale as either "clinically applicable" or "providing them with information they wanted to measure," depending on whether they had a clinical or a research perspective.

From an assessment standpoint, most respondents indicated a need to examine the relationship between fatigue and function—how people managed when fatigued, particularly in daily activities. Psychosocial and emotional factors were also considered essential components to assess, as well as sleep patterns, lifestyle, and exercise. Because no one set of measures covers all of these areas, a multipronged approach or some type of assessment battery may offer a solution.

After the completion of this preliminary work, a consensus conference was held to achieve consensus on the best approach to evaluation and outcome measures for MS patients in the areas of gait and fatigue.

Consensus Conference

The goals of the consensus conference were to 1) gain a better understanding of outcome measures being

***Table 2.* Gait and fatigue measures used by rehabilitation professionals and average usefulness ratings**

Construct	Measure	No. (%) of responses	Scale usefulness[a]	
			Mean/10	Range (1–10)
Gait	Timed 25-Foot Walk	8 (22.2)	7.4	6–8
	Observation	8 (22.2)	6.4	2–8
	Dynamic Gait Index	2 (5.6)	8	8
	Timed Up and Go test	2 (5.6)	8	8
	Tandem gait	1 (2.8)	8	8
	Functional Independence Measure	1 (2.8)	8	8
	12-item Multiple Sclerosis Walking Scale	1 (2.8)	NR	NA
Fatigue	Interview	4 (11.1)	5.75	2–8
	Fatigue Severity Scale	4 (11.1)	5.25	5–6
	Fatigue Impact Scale	3 (8.3)	8	7–9
	Modified Fatigue Impact Scale	2 (5.6)	5	5
	Self-rating scale	2 (5.6)	5	2–8
	Rate of Perceived Exertion	1 (2.8)	2	2
	Epsworth Sleepiness Scale	1 (2.8)	6	6
	Brief Fatigue Index	1 (2.8)	5	5

Abbreviations: NA, not applicable; NR, no response.
[a]Participants were asked to rate the scales subjectively on a scale from 1 to 10, with 10 indicating "superior."

used by rehabilitation professionals treating people with MS; 2) establish consensus on gait and fatigue outcome measures; and 3) establish consensus on required follow-up for transfer of this knowledge to rehabilitation professionals. As a group of MS clinicians and researchers, we hoped to create consistency and improve communication with regard to gait and fatigue outcome measures among practitioners across disciplines. An underlying premise was that consistency in outcome measures would facilitate both the interpretation of research data and its clinical application.

Who Participated?

A total of 24 health professionals with expertise in MS, listed at the end of this report, participated in the consensus conference. They represented a broad interdisciplinary mix, including neurologists, physiatrists, nurses, occupational therapists, physical therapists, and social workers. They were from diverse practice settings—including MS centers, rehabilitation centers, and private practice settings—and were widely distributed geographically throughout Canada and the United States.

Format

Initially, an overview was provided of issues pertinent to outcome measures in fatigue and gait, including psychometric considerations and factors related to mobility in MS. This was followed by small-group discussions of the issues and large-group consensus building.

Considerations for Outcome Measures: Psychometric Properties

An *outcome* is "a characteristic or construct that is expected to change owing to the provision of a strategy, intervention, or program."[4] Thus, an *outcome measure* is an instrument, scale, tool, or index that will show a change as a result of a treatment or strategy for which change is expected. It is essential, therefore, that the measure matches the purpose of the intervention being tested. Ensuring the credibility of outcome measures requires testing for relevant psychometric parameters. These include validity, reliability, responsiveness, sensitivity, and clinical utility. Examples of types and definitions of psychometric parameters relevant to fatigue and gait outcome measures used in MS are shown in Appendix 1.

Fatigue in MS

Fatigue: No Longer an Invisible Issue

Although fatigue is a real and persistent issue for MS patients and professionals, for many years it was an "invisible" problem. In the past 20 years, a dramatic shift has occurred from doing little to address fatigue to developing clinical practice guidelines[2] and conducting randomized controlled trials of medical, pharmacologic, and energy-conservation strategies.[5-7] Only in the past few years has fatigue in MS been assessed and treated in clinical rehabilitation. No consensus has emerged, however, on outcome measures or the use of measurement techniques to understand the impact or severity of fatigue in MS or the factors that contribute to fatigue.

Why Is It Important to Measure Fatigue in MS?

Measuring fatigue in patients with MS is important for several reasons. First, it helps to determine how fatigue is interfering with the individual's functioning. It is known that everyone, both patients and nonpatients, experiences fatigue and that fatigue follows a normal diurnal cycle. In people with MS, however, fatigue is often described as occurring more often than usual and in such a severe form that it profoundly affects daily life. Thus, screening for fatigue is necessary to understand the consequences of fatigue in the lives of people with MS. Second, measuring fatigue can help identify the factors that may be contributing to the problem. These factors may be diverse and either directly related to MS (primary MS fatigue) or not (nonprimary MS fatigue). Determining the appropriate course of intervention for each person requires understanding these individual factors. Third, if disabling fatigue is identified, appropriate intervention can be initiated; many of the factors that contribute to or exacerbate fatigue in MS patients are treatable. Fourth, the presence of fatigue should be taken into account in the treatment of other symptoms or problems, so that therapeutic regimens that may require expenditure of energy, such as exercise programs, stretching sessions, or cognitive retraining programs, can be tailored to maximize the patient's available energy.

Current Fatigue Measures in MS

Many different measures are used to evaluate fatigue in MS, but they inconsistently identify the potential impact and its severity. Measures of fatigue are generally subjective, often involve the use of paper-and-pencil or computer-input forms, and use a yes/no, Likert scale, or visual analogue scale (VAS) format. These measures are easy to score, use an ordinal or numerical scale, have between 1 and 40 items, and are usually ret-

rospective—that is, asking the patient about fatigue during a previous time period such as in the last day or week. For most measures, psychometric or clinicometric properties have been established.

Fatigue measures reported in the literature can be divided into two categories: general measures and population-specific measures. Examples of populations for which specific fatigue evaluations have been developed are people with cancer, HIV/AIDS, arthritis, brain injury, cardiac problems, Parkinson's disease, myasthenia gravis, and MS. Of fatigue measures used for patients with MS, five are MS-specific and five have been tested with MS as well as other patient populations.* These ten measures of fatigue fall into three formats: short, categorical, and noncategorical.

Short Measures

The short measures of fatigue are those that have less than ten items; they include the VAS for fatigue, the FSS, and the Daily Fatigue Impact Scale (D-FIS). The VAS, a single-item measure, has been shown to be a moderately reliable, valid, and useful tool for the rapid screening of fatigue impact.[8] The scale is represented by a 10-cm horizontal line, with one end representing "not fatigued at all" and the other end representing "extremely fatigued." The line is usually divided into thirds indicating 1) that fatigue is not a problem, 2) that there are periodic issues with fatigue, and 3) that fatigue is a substantial problem. A major advantage of this scale is that it takes very little time to administer.

The FSS is a nine-item measure that uses a Likert scale response. Patients rate statements on a scale of 1 to 7, with 1 indicating "strongly disagree" and 7 indicating "strongly agree." Higher scores indicate increased severity of fatigue. Many clinicians like this scale because it is quick and easy to administer. Since it is a global scale, however, it does not pinpoint the precise issue of concern related to fatigue.

The D-FIS is an eight-item measure that also uses a Likert scale format. It asks respondents to rate statements on a scale of 0 to 4, with 0 indicating "no problem" and 4 indicating "extreme problem." Higher scores indicate increased severity of fatigue. It should be noted that the metric testing for the D-FIS has been demonstrated only in Spanish.

Categorical Measures

Categorical measures of fatigue are those that have subscales or groupings of items; they include the FIS, MFIS, Multi-component Fatigue Scale (MFS), and Fatigue Descriptive Scale (FDS). The FIS, introduced in 1994, was the first scale to be described in the literature.[9] Originally, it contained 36 items. After additional research, the FIS was reduced to 21 items and renamed the MFIS. Both the FIS and MFIS use a Likert scale format and incorporate the three subscales of cognitive, physical, and social dimensions.

The MFS uses the Likert scale format and includes two subscales: cognitive fatigue and physical fatigue. The cognitive fatigue subscale includes items that differ slightly from those of other cognitive scales, and the physical fatigue scale includes questions related to sleep and drowsiness.

The FDS includes four sections: modality, severity, frequency, and Uhthoff's phenomenon (temperature-related issues). This evaluation uses a narrative or interview method for collecting and recording data. The FDS was developed by a physician and is usually administered only by physicians.

Noncategorical Measures

Noncategorical fatigue measures do not divide items into groupings or subscales but rather consist of a diverse set of items. These measures include the Wurzburg Fatigue Inventory for MS (WEIMus), Fatigue Assessment Instrument (FAI), and Rochester Fatigue Diary (RFD).

The WEIMus is a recently published scale that is currently being tested in English after being translated from German. This 17-item scale includes diverse aspects of fatigue and uses a Likert scale format.[10]

The FAI was originally developed for use in patients with cancer. It is a 29-item scale that has also been tested in people with MS.[11] Patients rate their responses to statements from 1 to 7, with 1 indicating "completely disagree" and 7 indicating "completely agree."

The RFD is a series of 24 VASs. Respondents mark a line on the chart hourly to indicate their energy level at that time. This is considered a temporal tool, as it is used throughout the day.

Each of these tests or measures seems to serve an important purpose. Our next task was to determine the

* More recently, two additional measurement systems have been developed: PROMIS (http://www.nihpromis.org/default.aspx) and NeuroQol (http://www.neuroqol.org/default.aspx). They are currently being tested with MS as well as other populations. They are computerized item-bank measurement systems that are psychometrically robust and may hold considerable promise.

optimal outcome measure for MS patients with fatigue. Fully understanding issues related to fatigue, however, requires both a global outcome measure and an assessment battery.

Reaching Agreement on Fatigue Measures

With a sense of the measures available, consensus conference participants discussed the characteristics of an optimal tool for measuring fatigue and related issues. The discussion focused on fatigue scales for clinical purposes rather than research purposes. Among the issues debated were fatigue impact versus fatigue severity, which are considered to be two different entities.

Conference participants agreed that measures must be responsive to change for diverse purposes, such as reimbursement, qualifying for disability benefits, insurance, and demonstrating treatment effect for medical, pharmaceutical, and behavioral interventions. It was agreed that it was ambitious to expect one measure to fulfill all of those purposes.

It was also agreed that fatigue outcome measures in the form of patient self-reports are needed and should include screens for a variety of contributing factors. Several questions were raised: Should there be a conceptual model? Should there be a framework that aligns with clinical processes? What do patients and clinicians need? Should outcome measures be clinician-based, self-administered, or both? If fatigue in MS results from numerous factors, how can they be accounted for in the fatigue evaluation?

A need was identified for both a fatigue *outcome measure* for MS that could show change and a fatigue *assessment battery* for MS that would screen for and identify factors contributing to fatigue, assist with informed clinical decision-making, and guide intervention decisions. Both would contribute significantly to clinical practice. For the purposes of this consensus conference, however, a fatigue outcome measure for MS became the focus.

Attention then turned to determining the characteristics of such a fatigue outcome measure. It was agreed that the measure should 1) include a screen for the functional ramifications of fatigue for activities and participation as defined in the World Health Organization's 2001 *International Classification of Functioning, Disability, and Health;* 2) be quick and easy to administer; 3) demonstrate psychometric integrity for MS; and 4) examine fatigue over a continuum of the MS disease course. This measure could be a self-report or clinician-administered. It was decided that the outcome measure should incorporate the elements of self-efficacy and a screening of past behaviors and methods used to deal with fatigue.

Although considerable debate ensued regarding the pros and cons of the various scales and their usefulness, a careful review of the scales in relation to these characteristics was required. A post–consensus conference analysis of the ten instruments currently used to measure fatigue in MS described above with reference to these characteristics and elements revealed that all of the measures require further development to fulfill these criteria (Table 3).

Mobility Issues in MS

The other aspects of functioning discussed during this CMSC consensus conference were ambulation and gait measures. As with fatigue, gait is difficult to define. The more it is discussed, the more elusive the definition seems to become.

The Conundrum of a Single Measure of Gait

Gait is a major area of activity limitation and participation restriction in MS. Clinicians and researchers must use accurate and quantifiable measures of gait performance in order to determine whether patients are progressing over time and how intervention(s) should be targeted (eg, focusing on vertical head turns or reducing multitasking). Good gait assessments should uncover other dysfunction that may not be immediately obvious.

An ideal gait measure would be quick and easy to administer, would not require major investments in training or materials to implement, and, most important, would be accurate. As with fatigue, psychometric properties are the key considerations in gait measures. Is the measure appropriate for people with MS? Are the results repeatable over time and between or among raters?

On the surface, developing a gait measure that incorporates these ideals does not seem difficult; further exploration, however, reveals several problems with this process. First, gait is not a unitary entity but encompasses many independent and interdependent variables—such as strength, motor control, range of motion, and sensation—that must be assessed separately and together. Depending on the test or gait tasks, gait will change, which suggests that a single measure

Table 3. Critique of commonly used fatigue outcome measures in MS

Fatigue outcome measure for clinical purposes	No. of items	Measures impact of fatigue on activities and participation	Time required (min)	Established metrics	Self- or clinician-administered	Items include		
						Self-efficacy	Managing fatigue	Contributing factors
MS only								
FIS (original)	36	Yes	20–30	Yes	Self	Yes	No	No
MFIS (modified)	21	Yes	15	Yes	Self	No	No	No
FDS	9	Yes	15	No	Clinician	No	No	No
RFD (diary)	24	No	10	Yes	Self	No	No	No
WEIMuS	17	Yes	10–15	In German	Self	No	No	Yes (1 item)
MS and other populations								
D-FIS (daily)	8	Yes - activity No - participation	5–10	In Spanish	Self	No	No	No
FSS	9	Yes	5–10	Yes	Self	No	No	No
VAS	1	No	1	Yes	Self	No	No	No
FAI	29	Yes	20–25	Yes	Self	No	Yes	Yes (5 items)
MFS	15	No	10–15	Yes	Self	No	No	No
PROMIS-Fsf	7	Yes	5	Yes	Self	No	No	No

Abbreviations: D-FIS, Daily Fatigue Impact Scale; FAI, Fatigue Assessment Instrument; FDS, Fatigue Descriptive Scale; FIS, Fatigue Impact Scale; FSS, Fatigue Severity Scale; MFIS, Modified Fatigue Impact Scale; MFS, Multi-component Fatigue Scale; MS, multiple sclerosis; PROMIS-Fsf, Patient Reported Outcome Measurement Information System–Fatigue short form; RFD, Rochester Fatigue Diary; VAS, visual analogue scale; WEIMuS, Wurzburg Fatigue Inventory for MS.

of gait will not suffice, as multiple aspects of gait can be measured. Among them are speed, distance, fatigue, gait deviations, force and angles, kinetic and kinematic changes, fall risk, and use of assistive devices.

Gait disorders in MS can present unique challenges. Unlike some other neurologic diseases, MS is highly heterogeneous. Because each patient has an individual gait pattern, there is no benchmark or "typical" MS gait, although this concept has been explored and some common characteristics have been presented.[12] This makes it difficult to identify a single "MS gait measure."

Gait disorders in MS can occur for many different reasons. Primary causes include spasticity, weakness, sensation loss through proprioception or kinesthesia, ataxia, diminished motor control, visual disturbance, fatigue, and vestibulopathy. Exacerbation of any of these symptoms may lead to gait impairment and fatigue. Secondary causes of gait impairment include contractures, disuse weakness due to lack of exercise, and secondary fatigue. Tertiary causes of gait impairment may relate to environmental factors, including walking surface, ambient temperature, and use of medications that affect gait, both directly and through fatigue-related side effects. Secondary and tertiary causes may compound the problem of repeatability of measuring the different aspects of gait such as distance and speed.

Gait in MS can be measured in multiple ways, and multiple potential deviations may occur. Thus, gait is an uncertain and variable construct with an uncertain and variable measure. The challenge of the conference was to develop a consensus on the optimal measures of gait outcomes within the diverse population of people with MS.

Current Gait Measures in MS

The survey identified the most common gait measures used in MS as the T25FW, observation, the DGI, the TUG test, the MSWS-12, and the Functional Independence Measure (FIM) (Table 4). Other measures discussed during the conference include the Expanded Disability Status Scale (EDSS), the Hauser Ambulation Index, the Tinetti Gait Assessment, the 6-Minute Walk, and kinetic and kinematic analysis (Table 5).

Many other tools are used in rehabilitation to examine balance and broader components of mobility, such as the Rivermead Mobility Index. In addition, some of the measures mentioned do not represent "true" gait or ambulation measures. These were nevertheless included to provide information on what is being used and highlight the difficulty of measuring

Table 4. Commonly used gait measures identified in survey of CMSC rehabilitation professionals

Gait measure	Characteristics	Advantages	Disadvantages	Psychometric properties
Timed 25-Foot Walk	Person is timed walking 25 feet, as fast as he/she is able safely	Easy to administer Inexpensive	Unable to assess gait deviations due to fatigue	High inter-rater and test-retest reliability and good concurrent validity[13,14]
Observation	Clinician observation of an individual's gait, usually in a controlled clinical setting	Easy to administer Inexpensive	Requires training to understand normal and abnormal gait	Poor inter-rater reliability
Dynamic Gait Index	8-point ordinal scale Score range 0–24 Score of <19 correlates to high falls risk	Easy to administer Inexpensive	Overlap of categories Does not assess fatigue due to short distances	Valid and reliable in people with MS[15]
Timed Up and Go test	From a seated position, an individual is asked to stand, walk 3 m, turn around, and walk back to the chair at a regular pace Score of ≤14 s has been shown to indicate high risk of falls	Easy to administer Inexpensive Correlates with gait speed and balance	Unable to assess gait deviations due to fatigue	Validity and reliability for fall prediction in community-dwelling older adults[16]
12-item Multiple Sclerosis Walking Scale	12-item self-report Score range 0–100 Higher scores reflect greater limitations in walking	Easy to administer Inexpensive Provides patient's perspective	Self-report questionnaire	Valid, reliable, and responsive patient-based measure of impact of MS on walking[17,18]
Functional Independence Measure	18-item, 7-level ordinal scale Completed by observation, in conference, or by telephone Two dimensions: motor and cognitive	Designed to assess dysfunction in individuals with progressive, fixed, or reversible neurologic disorders	Requires training Fee required	Valid and reliable tool for assessing people with MS[19,20]

Abbreviations: CMSC, Consortium of Multiple Sclerosis Centers; MS, multiple sclerosis.

gait. Measures of balance and broader components of mobility will be considered in a future forum, such as a consensus conference.

Because clinicians must evaluate a range of criteria to assess gait in MS patients, a single scale will not provide a comprehensive picture. A complete MS gait evaluation must include several relevant components, incorporating fatigue, speed, distance, balance and falls risk, observation, and use of assistive devices. Assessments should also be easy to administer, inexpensive, valid, and reliable, and should provide the clinician information that is useful in developing a treatment plan. Multiple scales—or an assessment battery—are necessary to provide a comprehensive view of the patient and incorporate all of the variables involved in measuring gait in people with MS.

Reaching Agreement on Gait Measures

The consensus conference revealed agreement, in principle, on the need for a means of measuring and assessing these variables. The general discussion revolved around the "what, why, and how" of determining appropriate measures and the best ways to encourage other rehabilitation professionals to adopt them.

Table 5. Commonly used gait and mobility measures used in MS but not identified through survey of CMSC rehabilitation professionals

Gait measure	Characteristics	Advantages	Disadvantages	Psychometric properties
6-Minute Walk[a]	Distance walked is measured over 6-min time period Walking is self-paced	Easy to administer Inexpensive Provides valuable information on effects of fatigue on ambulation	Time to administer Does not account for changes over the 6 min (eg, whether the first few minutes are faster than the last) Qualitative changes are not assessed	High inter-rater and intrarater reliability[21]
Expanded Disability Status Scale	20-point scale 0–10 with half-point increments (except with no 0.5 on scale) Scores of 0–4.5 indicate person is "fully ambulatory" for 50 m Scores of ≥5.0 indicate ambulatory deficits or other functional impairment	Easy to administer Widely used and tested	Ambulation defined only in terms of distance and assistance Qualitative changes are not assessed	Valid as a disability measure, but not specific to ambulation[22]
Tinetti Gait and Balance Assessment	Separate gait and balance scores Gait portion scored as 0–16 points Examines different phases of gait, identifying deficits or contralateral differences	Easy to administer Inexpensive Can be combined with balance portion to determine risk of falls	Has not been tested in MS for validity and reliability	Good to excellent inter-rater and intrarater reliability; mobility test scores highly correlated with gait speed in Parkinson's disease[23]
Hauser Ambulation Index	Ordinal scale combining qualitative, self-report, and observation 10 grades from 0 to 9, with 0 being normal and 9 being unable to transfer independently	Quick and easy to administer No specialized training required	Not responsive	Good inter-rater and test-retest reliability and convergent validity[24]
Kinetic and kinematic analysis	Measures force and angle of joints during gait cycle Provides data on spatial and temporal gait parameters	Provides precise, objective data	Expensive Requires special training Biomechanics may not reflect activity limitations/participation restrictions	Moderate reliability in pediatric population[25] No data found specific to MS population

Abbreviations: CMSC, Consortium of Multiple Sclerosis Centers; MS, multiple sclerosis.
[a]Some clinicians use a 2-min walk, but this has not been researched.

It is important to understand how these measures and variables behave in different settings, such as in-home, outpatient, and inpatient clinic settings. Measures should be continuous over a period of time and have strong psychometric properties. Space and safety must be considered, as well as variables such as time, temperature, and medication use. It was also acknowledged that the tools used by researchers and clinicians differ, and that the consensus conference would focus on the clinical setting. We also debated other issues,

such as deviations for higher-functioning patients, the pros and cons of various additional measures, and how many of these measures, in what combinations, should be used. Consideration must be given to insurance reimbursement, safety, and behavior modification, as well as the validity, reliability, and sensitivity of the measure or measures.

The most useful document that could result from this consensus conference would be a chart describing outcome measures for gait with information such as the type of tool (self-report vs clinician-administered); time, expense, and equipment required; and psychometric properties available. Most rehabilitation professionals want to know whether performance of a test is realistic in his or her specific setting; what equipment is needed; and the appropriate setting, such as neurology clinics or inpatient or outpatient rehabilitation clinics. A preliminary chart was established with characteristics for each measure (Table 6).

It was agreed that in a clinical setting, the T25FW, TUG test, DGI, 6-Minute Walk (and possibly a 2-minute version), and self-reported MSWS-12 should be the tools initially included in a gait outcome measures chart. Inclusion of psychometrics—such as validity, reliability, internal consistency, sensitivity, and responsiveness—is also extremely important. As noted in Table 5, however, much of this data is not currently available. Therefore, further research and development of this resource is needed. This chart provides a sense of what tools to use and can also answer the questions of when, why, and how.

Next Steps for Rehabilitation Outcomes

As previously noted, in addition to the chart for gait outcome measures, conference attendees also agreed on the need for a fatigue assessment battery. Of primary interest was information regarding specific and consistent outcome measures, to be disseminated to therapists throughout North America to improve decision-making. The goal is to establish clear communication, guidelines, and criteria for selecting the best outcome measures for both fatigue and gait.

We are pleased that we were able to reach consensus to move in this direction. Clearly, more work is needed. Many specific topics related to both fatigue and gait need to be addressed, and validating some of these measures in MS will be time-consuming. Conference participants have already begun work on both the fatigue assessment battery and a more detailed gait outcome measures chart. In researching existing tools and developing new ones, it is also important to consider the interaction between gait and fatigue. We believe that this consensus conference marks an important step

***Table 6.* Preliminary gait measures chart with measures identified at consensus conference**

Gait measure	Administration	Time and expense	Equipment needed	Psychometric properties
Timed 25-Foot Walk	Clinician-administered	Minimal time (1–2 min) No additional expense	Tape measure, stopwatch	Content and construct validity in MS population compared with EDSS[11]
Dynamic Gait Index	Clinician-administered Training recommended	Minimal time No additional expense	Stopwatch	Valid and reliable in MS population[12]
12-item Multiple Sclerosis Walking Scale	Self-report	Time varies for individuals, but should be ≤10 min	Survey, pen/pencil	Valid, reliable, and responsive patient-based measure of impact of MS on walking[14,15]
Timed Up and Go test	Clinician-administered	Minimal time (1–2 min) No additional expense	Armchair, tape measure, stopwatch	No validity and reliability data for MS population
6-Minute Walk	Clinician-administered Training recommended	6 min of testing time; need to provide time for adequate recovery	Stopwatch, premeasured distance	High inter-rater and intrarater reliability[17]

Abbreviations: EDSS, Expanded Disability Status Scale; MS, multiple sclerosis.

toward improving understanding of gait and fatigue among MS rehabilitation professionals. The conference participants identified a wide range of outcome measures for both gait and fatigue. The survey responses, as well as discussions among colleagues, make it clear that rehabilitation professionals are only moderately satisfied with these scales, and a new outcome measure or, more likely, a testing battery needs to be developed. We hope that this consensus conference lays the groundwork for creating more consistency in assessment and outcome measures and improving communication among practitioners and across disciplines. □

Consensus Conference Participants: Susan Bennett, Francois Bethoux, Theodore Brown, Kathleen Costello, Susan Forwell, Cindy Gackle, Eduard Gappmaier, Myrna Harden, Jodie Haselkorn, Jutta Hinrichs, Brian Hutchinson, Herb Karpatkin, Patricia Kennedy, Seema Khurana, Virgil Mathiowetz, Deborah Miller, Marie Namey, Melissa Pei, Patricia Provance, Tammy Roehrs, Anjali Shah, Christine Smith, Matthew Sutliff, Libby Winkler

Acknowledgments: The authors acknowledge the Foundation of the CMSC, which provided a grant in support of this consensus conference; and Lori Saslow, a professional medical writer funded by the CMSC, for her assistance with writing the manuscript.

Financial Disclosures: Before becoming an employee of Acorda Therapeutics, Mr. Hutchinson received consulting fees from Acorda. He has also received lecture fees from Medtronic Neurological, EMD Serono, and Teva Neuroscience. Ms. Bennett has received lecture fees from Biogen Idec, Acorda Therapeutics, Teva Neuroscience, and EMD Serono. Dr. Brown has received consulting fees from Bayer, EMD Serono, Acorda Therapeutics, and Teva Neuroscience; lecture fees from Bayer, Teva Neuroscience, and EMD Serono; and grant support from EMD Serono, Forest Labs, and Lilly. Dr. Forwell has received lecture fees from Teva Neuroscience and EMD Serono. Mr. Karpatkin and Dr. Miller have no conflicts of interest to disclose.

PracticePoints

- Rehabilitation outcome measures in MS are important to coordinate treatment, monitor progress toward established goals, improve quality of care, and maximize third-party reimbursement.
- Consensus on outcome measures can increase consistency in assessment and treatment among rehabilitation professionals.
- In people with MS, gait and fatigue are closely related and should be considered together.
- Additional research is needed to determine the most appropriate gait and fatigue outcome measures to use in clinical and rehabilitation settings.

Appendix 1. Psychometric characteristics of fatigue and gait outcome measures in multiple sclerosis

Reliability: consistency or repeatability of test scores

1. Inter-rater reliability: consistency between two or more different raters
2. Test-retest reliability (also known as stability): consistency of results over time
   - Meaningful when a parameter is expected to stay stable or a behavior is not expected to change over the time between test administrations, eg, 48 hours
3. Alternate form reliability: stability of test form, ie, altering sequencing of test items without a result change

Depending on the measure, the results of reliability testing can be expressed as:

a) percent agreement (ideal percent agreement is 100%; range, 0–100%)

b) intraclass correlation coefficient*

c) other variations of the correlation coefficient*

**Perfect reliability for all variations of correlation coefficient is 1; range, 0–1.*

Internal consistency (IC): consistency within the test, eliminating the influence of time

- Relevant only for tests with multiple items
- High IC suggests that the test as a whole measures one homogeneous construct, while low IC suggests that constructs being measured have heterogeneous factors.

Depending on the measure, the results of IC testing can be expressed as a Cronbach α or using the Kuder-Richardson formula.

Validity: appropriately, meaningfully, and usefully measures targeted behavior or entity and not something else

1. Content validity: How well the test samples the behavior of interest
   - "Face validity": On the surface, users, readers, and/or the person being tested agree that the test captures the behavior of interest. Results are built into item generation and instrument development.
2. Criterion validity: How well the test captures the behavior of interest relative to an external criterion, usually a "gold standard." Results are expressed as a correlation coefficient.
3. Construct validity: Test correlates with variables to which it should be theoretically related or the process of testing hypotheses about the behavior in question.

Results can be based on searching for alternative explanations that might account for the observed performance or a statistical factor analysis of which there are several types, eg, principal components analysis or cluster analysis.

Responsiveness: ability of the test to detect change, when change in fact has occurred

Sensitivity: ability of the test to correctly classify a condition, performance, or situation

Specificity: ability of the test to correctly classify those not having a condition or performance issue

References

1. Bennett S, Thrower B, Foley F, et al. Developing a consensus statement for rehabilitation outcome measures in multiple sclerosis. Poster presented at: 18th Annual Meeting of the Consortium of Multiple Sclerosis Centers; June 2004; Toronto, Ontario, Canada.
2. Multiple Sclerosis Council for Clinical Practice Guidelines. *Fatigue and Multiple Sclerosis: Evidence-Based Management Strategies for Fatigue in Multiple Sclerosis*. Washington, DC: Paralyzed Veterans of America; 1998.
3. Forwell S, Brunham S, Tremlett H, Morrison W, Oger J. Primary and nonprimary fatigue in multiple sclerosis. *Int J MS Care.* 2008;10: 14–20.
4. Finch E, Brooks D, Stratford P, Mayo N. *Physical Rehabilitation Outcome Measures: A Guide to Enhanced Decision Making*. 2nd ed. Hamilton, Ontario: BC Decker; 2002.
5. Sheean GL, Murray NM, Rothwell JC, Miller DH, Thompson AJ. An open-labelled clinical and electrophysiological study of 3,4 diaminopyridine in the treatment of fatigue in multiple sclerosis. *Brain*. 1998;121:967–975.
6. Rammohan K, Rosenberg J, Lynn D, Blumenfeld A, Pollak C, Nagaraja H. Efficacy and safety of modafinil (Provigil®) for the treatment of fatigue in multiple sclerosis: a two centre phase 2 study. *J Neurol Neurosurg Psychiatry.* 2002;72:179–183.
7. Mathiowetz VG, Finlayson ML, Matuska KM, Chen HY, Luo P. Randomized controlled trial of an energy conservation course for persons with multiple sclerosis. *Mult Scler.* 2005;11:592–601.
8. Kos D, Nagels G, D'Hooghe MB, Duportail M, Kerckhofs E. A rapid screening tool for fatigue impact in multiple sclerosis. *BMC Neurol.* 2006;6:1–8.
9. Fisk JD, Ritvo PG, Ross L, Haase DA, Marrie TJ, Schlech WF. Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. *Clin Infect Dis*. 1994;18(Suppl 1):S79–83.
10. Flachenecker P, Müller G, König H, Meissner H, Toyka KV, Rieckmann P. "Fatigue" in multiple sclerosis: development and validation of the "Wurzburger Fatigue Inventory for MS" [in German]. *Nervenarzt.* 2006;77:165–166, 168–170, 172–174.
11. Schwartz JE, Jandorf L, Krupp LB. The measurement of fatigue: a new instrument. *J Psychosom Res*. 1993;37:753–762.
12. Martin CL, Phillips BA, Kilpatrick TJ, et al. Gait and balance impairment in early multiple sclerosis in the absence of clinical disability. *Mult Scler.* 2006;12:620–628.
13. Rudick RA, Cutter G, Reingold S. The Multiple Sclerosis Functional Composite: a new clinical outcome measure for multiple sclerosis trials. *Mult Scler.* 2002;8:359–365.
14. Kaufman M, Moyer D, Norton J. The significant change for the Timed 25-foot Walk in the multiple sclerosis functional composite. *Mult Scler.* 2000;6:286–290.
15. McConvey J, Bennett SE. Reliability of the Dynamic Gait Index in individuals with multiple sclerosis. *Arch Phys Med Rehabil.* 2005; 86:130–133.
16. Shumway-Cook A, Brauer S, Wollacott M. Predicting the probability of falls in community-dwelling older adults using the Timed Up & Go Test. *Phys Ther.* 2000;80:896–903.
17. Hobart JC, Riazza A, Lamping DL, Fitzpatrick R, Thompson AJ. Measuring the impact of MS on walking ability: the 12-Item MS Walking Scale (MSWS-12). *Neurology.* 2003;60:31–36.
18. McGuigan C, Hutchinson M. Confirming the validity and responsiveness of the Multiple Sclerosis Walking Scale-12 (MSWS-12). *Neurology.* 2004;62:2103–2105.
19. Brousseau L, Wolfson C. The inter-rater reliability and construct validity of the Functional Independence Measure for multiple sclerosis subjects. *Clin Rehabil.* 1994;8:107–115.
20. Granger CV, Fielder RC. The reliability of the FIM: a quantitative review. *Arch Phys Med Rehabil.* 1996;77:1226–1232.
21. Goldman M, Marrie RA, Cohen J. Evaluation of the six-minute walk in multiple sclerosis subjects and healthy controls. *Mult Scler.* 2008;14:383–390.
22. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology.* 1983;33: 1444–1452.
23. Kegelmeyer DA, Kloos AD, Thomas KM, Kostyk SK. Reliability and validity of the Tinetti Mobility Test for individuals with Parkinson disease. *Phys Ther.* 2007;87:1369–1378.
24. Hauser SL, Dawson DM, Lehrich JR, et al. Intensive immunosuppression in progressive multiple sclerosis: a randomized three arm study of high dose intravenous, cyclophosphamide, plasma exchange, and ACTH. *N Engl J Med.* 1983;308:173–180.
25. Krebs DE, Edelstein JE, Fishman S. Reliability of observational kinematic gait analysis. *Phys Ther.* 1985;65:1027–1033.