

Appendix 1 (as supplied by authors): Psychometric properties of outcome instruments

Primary outcome

Modified Ashworth Scale

The literature on inter-rater reliability and test-retest reliability of the Ashworth Scale and Modified Ashworth Scale provides mixed reviews. While several studies demonstrate good inter-rater reliability of the Modified Ashworth Scale (Bohannon et al., 1987; Bodin et al., 1991; Sloan et al., 1992; Allison et al., 1996), others have questioned the reliability of the measure (Pandyan et al., 1999; Nuyens et al., 1994). The Ashworth and Modified Ashworth Scales represent the most widely used and accepted measures of spasticity in clinical practice.

Allison SC, Abraham LD, Petersen CL. Reliability of the modified Ashworth Scale in the assessment of plantar flexor muscle spasticity in patients with traumatic brain injury. Int J Rehabil Res 1996; 19:67–78.

Bodin PG, Morris ME. Inter rater reliability of the modified Ashworth Scale for wrist flexors spasticity following stroke. World Federation of Physiotherapy, 11th Congress 1991: 505–507.

Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. Phys Ther. 1987;67:206–207.

Nuyens G, De Weerdts W, Ketelaer P. Interrater reliability of the Ashworth Scale in multiple sclerosis. Clin Rehabil. 1994;8:286–292.

Pandyan AD, Johnson GR, Price CIM, et al. A review of the properties and limitations of the Ashworth and the modified Ashworth scales as measures of spasticity. Clinical Rehabilitation. 1999;13:373–383.

Sloan RL, Sinclair E, Thompson J, Taylor S, Pentland B. Inter-rater reliability of the modified Ashworth Scale for spasticity in hemiplegic patients. Int J Rehabil Res 1992; 15: 158–61.

Secondary outcomes

Visual Analogue Scale of Pain

Assesses current level of pain using a horizontal 100 mm line, anchored at one end by the descriptor “No Pain” and at the other by “Worst Pain Imaginable.”

The VAS has a number of validation studies, relating it to verbally-reported change in pain. The test-retest reliability of the VAS is high (e.g., ICC = .99 for 0 and 1 minute VAS scores). The minimum clinically significant difference in VAS pain scores (in relation to a verbal report of “a little better” or “a little worse”) has been estimated to be 9 points (95% CI, 6 to 13) for acute pain in an emergency department, while another study estimated it at 16 points, also in an acute ED context. *Gallagher EJ, Bijur PE, Latimer C, Silver W. Reliability and validity of a visual analog scale for acute abdominal pain in the ED. Am J Emerg Med. 2002 Jul;20(4):287-90;*

Kelly AM. Does the clinically significant difference in visual analog scale pain scores vary with gender, age, or cause of pain? Acad Emerg Med. 1998 Nov;5(11):1086-90.

Timed Walk

This is a mobility and leg function performance test in which the patient is timed walking 25 feet. It is similar to the Ambulation Index, differing only with regard to scoring and is strongly recommended for clinical trials to examine ambulation and leg function.

The Timed 25-Foot Walk is a component of the MS Functional Composite, which was developed by the National MS Society's Clinical Outcomes Assessment Task Force to be a multidimensional quantitative clinical outcome measure that includes tests of leg function/ambulation (Timed 25-Foot Walk), arm function (Nine-Hole Peg Test), and cognitive function (Paced Auditory Serial Addition Test) (Fischer 1999). The Timed 25-Foot Walk has been shown to have good predictive value for the Expanded Disability Status Scale (EDSS) (Rudick 2001), a clinical measure of severity of disability, and high test-retest reliability in individuals with lower EDSS scores (Cutter 1999). Cohen, et al found excellent intrarater and interrater reliability for the MSFC, of which the Timed 25-Foot Walk is a principle component. The intraclass correlation coefficient (ICC) for intrarater reliability was 0.97, and the ICC for interrater reliability was 0.95 and 0.96 when patients were reassessed 6 months later (Cohen 2000).

Cohen JA, Fischer JS, Bolibrush DM, Jak AJ, Kniker JE, Mertz LA, Skaramagas TT, Cutter GR. Intrarater and interrater reliability of the MS functional composite outcome measure. Neurology. 2000 Feb

Appendix to: *Corey-Bloom J, Wolfson T, Gamst A, et al. Smoked cannabis for spasticity in multiple sclerosis: a randomized, placebo-controlled trial. CMAJ 2012. DOI:10.1503/cmaj110837.*

Copyright © 2012 Canadian Medical Association or its licensors

22;54(4):802-6. PMID: 10690966 [PubMed - indexed for MEDLINE]

Cutter GR, Baier ML, Rudick RA, Cookfair DL, Fischer JS, Petkau J, Syndulko K, Weinshenker BG, Antel JP, Confavreux C, Ellison GW, Lublin F, Miller AE, Rao SM, Reingold S, Thompson A, Willoughby E. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain*. 1999 May;122 (Pt 5):871-82. PMID: 10355672 [PubMed – indexed for MEDLINE]

Fischer JS, Rudick RA, Cutter GR, Reingold SC. The Multiple Sclerosis Functional Composite Measure (MSFC): an integrated approach to MS clinical outcome assessment. National MS Society Clinical Outcomes Assessment Task Force. *Mult Scler*. 1999 Aug;5(4):244-50. Review. PMID: 10467383 [PubMed - indexed for MEDLINE]

Rudick RA, Cutter G, Baier M, Fisher E, Dougherty D, Weinstock-Guttman B, Mass MK, Miller D, Simonian NA. Use of the Multiple Sclerosis Functional Composite to predict disability in relapsing MS. *Neurology*. 2001 May 22;56(10):1324-30. PMID: 11376182 [PubMed - indexed for MEDLINE]

Paced Auditory Serial Addition Task (PASAT)

The PASAT has been recommended as a core measure in clinical trials involving MS patients. Test-retest reliability for the PASAT is generally high (.90-.97) and previous studies have demonstrated its validity in discriminating MS patients from controls. It is also sensitive to treatment effects in MS, and correlated with total lesion volume. Although there is no established minimally clinically important difference, a drop of 20 points in a cohort with the characteristics of the current study cohort would move the person from the normal to impaired range. *Tombaugh TN. A comprehensive review of the Paced Auditory Serial Addition Test (PASAT). Arch Clin Neuropsychol. 2006 Jan;21(1):53-76. Epub 2005 Nov 14; Fisk JD, Archibald CJ. Limitations of the Paced Auditory Serial Addition Test as a measure of working memory in patients with multiple sclerosis. J Int Neuropsychol Soc. 2001 Mar;7(3):363-72; Cohen, et al. Benefit of interferon B-1a on MSFC progression in secondary progressive MS. Neurology 2002, 59, 679-687; Hohol et al. Serial neuropsychological assessment and magnetic resonance imaging analysis in multiple sclerosis. Arch Neurol, 54, 1018-1025. Heaton, R.K., et al., Revised Comprehensive Norms for an Expanded Halstead Reitan Battery: Demographically Adjusted Neuropsychological Norms for African American and Caucasian Adults 2004, Lutz, FL: Psychological Assessment Resources.*

Brief Symptom Inventory

Assesses for the emergence of neuropsychiatric symptoms. A 53-item self-report checklist; each item is a 5-point scale of distress, ranging from “not-at-all” to “extremely”.

Cronbach's alpha for internal consistency on the nine subscales of the BSI ranges from .71-.85 and test-reliability ranges from .68-.91. The global indices from the BSI had test-retest reliability coefficients of .80-.90. The BSI has also been shown to have good convergent validity with the Minnesota Multiphasic Personality Inventory (MMPI). *Derogatis, L. R.: The Brief Symptom Inventory (BSI) Administration, Scoring, and Procedures Manual. (Baltimore, MD: John Hopkins University School of Medicine), 1982.*

Perceived Deficits Questionnaire

A self-report measure of cognitive dysfunction.

The PDQ has been used previously in MS to show that self-perceived cognitive dysfunction relates more to depression than objective cognitive dysfunction. The PDQ has been shown to have good reliability and validity in persons with MS (Cronbach's alpha .77-.97). *Lovera J, Bagert B, Smoot KH, Wild K, Frank R, Bogardus K, Oken BS, Whitham RH, Bourdette DN. Correlations of Perceived Deficits Questionnaire of Multiple Sclerosis Quality of Life Inventory with Beck Depression Inventory and neuropsychological tests. J Rehabil Res Dev. 2006 Jan-Feb;43(1):73-82; Fischer, J.S., LaRocca, N.G., Miller, D.M., Ritvo, P.G., Andrews, H., & Paty, D. (1999). Recent developments in the assessment of quality of life in multiple sclerosis (MS). Multiple Sclerosis, 5, 251–259; Marrie, R.A., Miller, D.M., Chelune, G.J., & Cohen, J.A. (2003). Validity and reliability of the MSQLI in cognitively impaired patients with multiple sclerosis. Multiple Sclerosis, 9, 621–626.*

Appendix to: Corey-Bloom J, Wolfson T, Gamst A, et al. Smoked cannabis for spasticity in multiple sclerosis: a randomized, placebo-controlled trial. *CMAJ* 2012. DOI:10.1503/cmaj110837.
Copyright © 2012 Canadian Medical Association or its licensors

Modified Fatigue Impact Scale

Examines the effects of fatigue in terms of physical, cognitive, and psychosocial functioning.

The mFIS has shown good correlation ($r = 0.68$, $p < 0.0001$) with the Fatigue Severity Scale (FSS), another common rating scale of fatigue, and with depression scores from the Beck Depression Inventory (BDI). Test-retest reliability for the mFIS was also high (ICC 0.84-0.93). Téllez N, Ríó J, Tintoré M, Nos C, Galán I, Montalban X. *Does the Modified Fatigue Impact Scale offer a more comprehensive assessment of fatigue in MS? Mult Scler. 2005 Apr;11(2):198-202*; Kos D, Kerckhofs E, Carrea I, Verza R, Ramos M, Jansa J. *Evaluation of the Modified Fatigue Impact Scale in four different European countries. Mult Scler. 2005 Feb;11(1):76-80.*

Subjective Ratings of High and Sedation, question 1

“How high do you feel?” (scale of 0-10), assessed after inhaled product on each day of each phase.

The SRHS-R was adapted from Block, et al (1998), who had originally adapted it from Norris (1971), for the present study. Block et al. demonstrated that the questionnaire was useful in dissociating various effects of cannabis. We added questions to assess patient motivation and perceived performance on study assessments. Block RI, Erwin WJ, Farinpour R, Braverman K (1998). *Sedative, stimulant, and other subjective effects of marijuana: relationships to smoking techniques. Pharmacol Biochem Behav 59:405–412.* Norris, H.: *The action of sedatives on brain stem oculomotor systems in man. Neuropharmacology 10:181–191; 1971.*