Web Supplement: 1 - Description of outcome measures

Physical activity

Godin Leisure-Time Exercise Questionnaire (GLTEQ)[1,2]

This is a self-report measure of usual physical activity that has been widely used in MS research. It consists of 2 items: the first measures frequency and intensity of exercise during free time in a typical week. These weekly frequencies are multiplied by metabolic equivalents and summed to form a measure of total leisure activity. The second question asks about the frequency of engaging in any regular activity long enough to work up a sweat with three response options provided (often, sometimes, never/rarely).

activPAL3 tri-axial accelerometer[3,4] **[http://www.paltechnologies.com]**

This classifies an individual's free-living activity into periods spent sitting, standing and walking and gives the number of steps and sit-to-stand episodes. Default activPAL settings were used and activPAL accelerometers were administered with instructions to wear for 24 hours/day for 7 consecutive days including at night but to remove during water-based activities (showering, bathing, swimming etc.). activPALs were affixed with 'PAL stickies' plus Hypafix medical tape if needed. An activPAL pack (including instructions with illustrations and a pre-paid envelope for postal return) was given to participants at the physical assessment appointment. The study coordinator was available to give a demonstration of how to affix the activPAL and could be contacted by telephone to answer any questions related to the activPAL. Participants were instructed to affix the activPAL to the mid-line of the thigh (a diagram was provided for guidance). We did not specify which leg as people could have had weakness in one leg. Participants were supplied with three additional PAL stickies and 10 strips of Hypafix medical tape. Instructions noted that the

activPAL should not be applied to broken skin and use should be discontinued immediately if an adverse skin reaction (such as a rash) or any discomfort was experienced.

Psychological well-being, quality of life

Hospital Anxiety and Depression Scale (HADS)[5]

The HADS is a self-report measure consisting of an anxiety and a depression subscale. Each subscale consists of 7 items with a 4-point Likert-type response scale. Higher scores indicate greater levels of anxiety and depression.

EuroQol-5 Dimensions-5 Levels (EQ-5D-5L)[6]

This is a measure of health status consisting of a descriptive system (five dimensions: mobility, self-care, usual activities, pain/discomfort, anxiety/depression). Each dimension has five levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. Additionally, a visual analogue scale records the respondent's self-rated health on a scale with endpoints labelled 'the best health you can imagine' and 'the worst health you can imagine'. Responses on the EQ-5D-5L can be converted to utility values using a value set for England.[7]

Multiple Sclerosis Impact Scale (MSIS-29)[8]

This scale measures the physical (20 items) and psychological impact (9 items) of MS on day-to-day life. It uses 5-point Likert-type scales ranging from 'not at all' to 'extremely' and is based on quality of life in the last two weeks.

The Fatigue Symptom Inventory (FSI)[9]

The FSI is a 14-item self-administered multi-dimensional questionnaire which measures the

severity (4 items - most, least, 'average, fatigue in past week and current fatigue), frequency in past week (2 items - how many days and how much of the day in past week), diurnal variation of fatigue (1 item) and its perceived interference on quality of life (7 items - general level of activity, ability to bathe and dress, normal work activity, ability to concentrate, relations with others, enjoyment of life, and mood).

The Medical Outcomes Short-Form Survey - version 2 (SF36 v. 2)[10]

The SF-36 measures eight dimensions of quality of life: physical functioning, role limitations because of physical health problems, bodily pain, general health perceptions, vitality, social functioning, role limitations because of emotional problems, and mental health. It uses Likert-type response scales and generates scores for the eight dimensions as well as two component summary measures (physical and mental) which we report in the current paper. SF-6D utility scores can be derived from scores on the SF-36.[11]

Self-efficacy

The Spinal Cord Injury Exercise Self-Efficacy Scale (SCI-ESES)[12]

This scale was originally developed and validated in spinal cord populations. However, it has since been used in MS.[13] It consists of 10 items with higher scores indicating higher perceived self-efficacy. It uses a 4-point Likert-type response scale.

The Multiple Sclerosis Self-Efficacy (MSSE) Scale[14]

This 18-item scale comprises two subscales: i. confidence with function (9 items) and ii. confidence with ability to manage symptoms/cope with the demands of illness (9 items). In the current study we only used the latter subscale. The rating scale consists of 10-points where 10 is 'very uncertain', 50 is 'moderately certain' and 100 is 'very certain'. Higher scores indicate greater self-efficacy.

Balance/gait/mobility

Two-Minute Walk Test[15]

This is a measurement of endurance by assessing walking distance over 2 minutes.
The individual is instructed to walk for two minutes and to cover as much ground as possible in that time. The distance covered is measured and logged. Habitual assistive devices and orthotics can be used.

Step Test[16]

This is a test of dynamic standing balance. It involves recording the number of times a participant steps one foot fully on, then off a block as quickly as possible in a 15 second time period. Two block heights are used: 7.5 cm and 15cm. Each leg is tested separately. Participants are instructed to perform this task as quickly as possible.

Steady Stance Test[17]

The steady stance test measures the ability of an individual to maintain a steady stance for a 60 second period in five predetermined stances without support: (feet apart, feet together, stride stance, tandem stance, single leg stance). The tests end when an individual either maintains steady stance for the 60 second testing duration, loses balances and takes support or alters their foot position.

Instrumented Timed Up and Go (i-TUG)[18]

This is a timed test used to examine functional mobility and requires the individual to stand up from a seated position, walk three metres, turn, walk back, and sit down. The accelerometer and telemetry system were attached to the participant's back with a strap.

Three-axis accelerometer data were recorded in real-time on a PC via telemetry. Time to undertake the test was derived from the accelerometer data. In the current study participants undertook one practice run followed by two recorded performances of the i-TUG test. Mean performance across the two recorded timed tests was used in the analysis.

Gait Stride-time Rhythmicity[19]

This was measured using using a portable recorder connected to flat in-shoe heel impact sensors. Participants were asked to walk unaccompanied on an extended walk in a non-laboratory setting along a covered and flat route (between 200-250 steps per foot). They were instructed to walk at their normal pace without interruption and could use their normal orthotic or walking aid. The stride time between adjacent heel strikes was recorded for each foot separately and the mean and standard deviation stride times were calculated.

Static Posturography[20]

The Poole Hospital Static Posturography System provided a method for objectively assessing balance via the tracking of limits of sway in a series of standardised conditions. It used ultrasound (transmitter attached to the participant's waist) time-of-flight (ToF) posturography measurements to locate a participant's centre of gravity (CoG) in a protocol mirroring the Equitest sensory organisation test protocol.[21] Anterior/posterior and lateral motions were detected independently by measuring variations in the ToF from the transmitters to appropriately positioned wall-mounted receivers. The motion of the CoG was accurately tracked during the course of 20 second assessment periods as the participant stood on either a solid or soft surface with eyes open or eyes closed or when doing a cognitive distractor task. We report the Equilibrium Quotient (EQ) percent score (100- (Anterior posterior maximum

sway during 20 seconds/maximum anterior posterior movement possible without losing balance) *100).

Hand dexterity/coordination

Nine-Hole Peg Test[22]

This timed test assesses finger and hand dexterity in both hands. It involves placing nine pegs into a peg board and then removing them as quickly as possible.

References

1. Godin G, Shephard RJ. A simple method to assess exercise behaviour in the community. Can J Appl Sport Sci 1985;10:141–146.
2. Gosney JL, Scott JA, Snook EM, *et al.* Physical activity and multiple sclerosis. Validity of self-report and objective measures. Fam Comm Health 2007;30:144–50.
3. Ryan CG, Grant PM, Tigbe WW, *et al.* The validity and reliability of a novel activity monitor as a measure of walking. Br J Sports Med 2006;40:779–7844.
4. Edwardson CL, Winkler EAH, Bodicoat DH, *et al.* Considerations when using the activPAL Considerations when using the activPAL monitor in field based research with adult populations. J Sport Health Sci 2016; doi: 10.1016/j.jshs.2016.02.002.
5. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand 1983;67:361–70.
6. The EuroQol Group. EuroQol – a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.
7. Devlin N, Shah K, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. Office of Health Economics: London, UK; 2016.
8. Hobart J, Lamping D, Riazi A. The multiple sclerosis impact scale (MSIS-29): a new patient-based outcome measure. Brain 2001;124:962–73.
9. Hann DM, Jacobsen PB, Azzarello LM, *et al.* Measurement of fatigue in cancer patients: development and validation of the Fatigue Symptom Inventory. Qual Life Res 1998;74:301–10.

10. Ware JE, Sherbourne CD. The MOS 36-item Short-Form health survey (SF-36): I. Conceptual framework and item selection. Med Care 1992;30:473–83.
11. Brazier JE, Roberts JR. The estimation of a preference-based index from the SF-12. Med Care 2004;42:851-859.
12. Kroll T, Kehn M, Ho PS, *et al.* The SCI Exercise Self-Efficacy Scale (ESES): development and psychometric properties. Int J Behav Nutr Phys Act 2007;4:34.
13. Stroud N, Minahan C, Sabapathy S. The perceived benefits and barriers to exercise participation in persons with multiple sclerosis.Disabil Rehabil 2009;31:2216–22.
14. Schwartz CE, Coulthard-Morris L, Zeng Q, *et al.* Measuring self-efficacy in people with multiple sclerosis: A validation study. Arch Phys Med Rehabil 1996;77:394–8.
15. Rossier P, Wade DT. Validity and reliability comparison of 4 mobility measures in patients presenting with neurologic impairment. Arch Phys Med Rehabil 2001;82:9–13.
16. Hill KD, Bernhardt J, McGann AM, *et al.* A new test of dynamic standing balance for stroke patients: reliability, validity and comparison with healthy elderly. Physiother Can 1996;48:257–62.
17. Frzoviv D, Morris ME, Vowels L. Clinical tests of standing balance; performance of persons with multiple sclerosis. Arch Phys Med 2000;81:215–21.
18. Cattaneo D, Regola A, Meotti M. Validity of six balance disorders scales in persons with multiple sclerosis. Disabil Rehabil 2006;28:789–95.
19. Perring S, Summers T. Laboratory-free measurement of gait rhythmicity in the assessment of the degree of impairment and the effectiveness of rehabilitation in patients with vertigo resulting from vestibular hypofunction. Physiol Meas 2007;28:697–705.
20. Perring S, Jepson R. The use of a two-dimensional ultrasound time-of-flight system for location of the centre of gravity in posturography measurements. Physiol Meas 2001;22:447–52.
21. EquiTest Operator's Manual, version 4.02, Appendix F: Technical Information, NeuroCom International, Clackamas, Oregon, 1992.
22. Mathiowetz V, Weber K, Kashman N, *et al.* Adult norms for the nine-hole peg test of finger dexterity. Occup Ther J Res 1985;5:24–37.