Journal of the Neurological Sciences 347 (2014) 268–274

Contents lists available at ScienceDirect

Journal of the Neurological Sciences

journal homepage: www.elsevier.com/locate/jns

Investigating the minimal important difference in ambulation in multiple sclerosis: A disconnect between performance-based and patient-reported outcomes?

Carolyn E. Schwartz [a,b,c,*], Armon Ayandeh [a], Robert W. Motl [d]

[a] DeltaQuest Foundation, Inc., Concord, MA, USA
[b] Department of Medicine, Tufts University Medical School, Boston, MA, USA
[c] Department of Orthopaedic Surgery, Tufts University Medical School, Boston, MA, USA
[d] Department of Kinesiology and Community Health, University of Illinois at Urbana-Champaign, Urbana, IL, USA

ARTICLE INFO

Article history:
Received 22 May 2014
Received in revised form 2 September 2014
Accepted 13 October 2014
Available online 18 October 2014

Keywords:
Multiple sclerosis
Ambulation
Patient-reported outcomes
Performance measure
Longitudinal construct validity

ABSTRACT

Objective: We sought to estimate the MID on two patient-reported outcome (PRO) measures that are frequently used in multiple sclerosis (MS) clinical research: the MS Walking Scale and the MS Impact Scale-29. We anchored the Minimally Important Differences with an objective measure of ambulation, the accelerometer.
Methods: This secondary analysis used longitudinal data from an observational study of symptoms and physical activity in 269 people with Relapsing–Remitting Multiple Sclerosis. Participants completed a battery of PRO questionnaires, and then wore an accelerometer for seven days at each data collection time point every six months for 2.5 years. Statistical analysis first defined Change Groups on the basis of the performance-based accelerometer scores, anchored to 0.5 standard deviation change; then change was defined on the basis of published and linked MIDs for the PROs.
Results: The performance-based (accelerometer) and PRO-based change distributions were stable over time. Raw scores among the accelerometer and PRO measures were associated with large effect sizes, and PRO change scores were associated with each other but not with accelerometer change scores.
Conclusions: These findings contradict a central assumption that may underlie clinical research studies: that a cross-sectional correlation implies that change in PROs will correspond with change in behavior/performance. Possible explanations related to accuracy of the performance-based measure, as well as response shift effects on the PROs are discussed.

1. Introduction

The use of patient-reported outcomes (PROs) in medical outcome research has grown in prominence and sophistication in the past two decades. Increasingly recognized as a source of important information that is not redundant with information reported by clinicians [1] or family-member caregivers [2], PROs provide the patient's perspective on symptom experience, symptom impact, and quality of life. Often using evaluative measurement tools which emphasize the subjective and idiographic nature of the variable human experience of health and illness, PRO tools face an increasingly rigorous validation process that characterizes and quantifies their reliability, validity, and responsiveness [3,4]. Technological advances in statistical software have facilitated these psychometric analyses, enabling the implementation of both classical- and item-response theory-based analyses that quantify aspects of reliability, validity and responsiveness in highly specific ways [5,6].

With this growth in technological prowess, the field of PRO research has developed thoughtful methods for evaluating the responsiveness of measurement tools to facilitate the interpretation of these measures [7]. Responsiveness is a key aspect of validity and recent guidelines for assessing responsiveness are useful in distinguishing types of responsiveness and how to evaluate it [8]. This growing research base on responsiveness has suggested that responsiveness is a highly contextual characteristic, affected by *who* is being measured for *what* outcomes in what research or clinical context (*where*) using what mode of data collection (*how*) and at what stage of the disease trajectory (*when*) [9]. Work has focused on understanding how much change is large enough to be discernible and regarded as important [10]. Referred to as the *Minimally Important Difference* (MID), this has been defined as "the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects

* Corresponding author at: DeltaQuest Foundation, 31 Mitchell Road, Concord, MA 01742, USA. Tel.: +1 978 758 1553.
E-mail address: carolyn.schwartz@deltaquest.org (C.E. Schwartz).

http://dx.doi.org/10.1016/j.jns.2014.10.021
0022-510X/

and excessive cost, a change in the patient's management" [11]. The MID may be estimated by taking an initial or baseline assessment and a follow-up assessment, and at follow-up asking the patient how much their condition has changed (i.e., a transition rating or global rating of change) [10]. Using this transition rating as an anchor, one can estimate the mean change in the assessment that corresponds to getting worse or getting better. The methodological challenge of using such patient-reported transition ratings is the potential biases due to response shift, recall bias, and implicit theories of change [12–14].

These potential biases have perhaps alerted investigators to examine the consistency of MIDs across studies and to note variability and inconsistency in meaningful-change metrics. Even in measures of relatively concrete behaviors, such as ambulation, there seems to be variability in the amount of change that corresponds to a person's impression of clinically-important change [15]. For example, past research on MID of the Multiple Sclerosis Walking Scale-12 (MSWS) [16] has yielded varying MID estimates, ranging from 4 to 10 points on a 100-point scale [15,17,18]. Differences between patient groups or studies in what constitutes an important change could impair the comparability of PRO data on the same instrument(s) across studies [19].

In response to the challenge of 'moving goal posts' [20,21], we sought to estimate the MID on two PRO measures that are frequently used in multiple sclerosis (MS) clinical research: the MSWS [16] and the MS Impact Scale-29 (MSIS) [22]. We anchored the MIDs with an objective measure of ambulation, the accelerometer [23]. We used the well-documented robustness of the half-standard deviation of the accelerometer change score as a benchmark for clinically important change [24] to estimate the MID of the MSWS and the MSIS. We then investigated relationships between accelerometer change and PRO change over time, and examined self-efficacy as a psychosocial factor that may explain discrepancies between objective and patient-reported change.

2. Methods

2.1. Sample

This secondary analysis used data from an observational study of symptoms and physical activity over 2.5 years in people with Relapsing–Remitting Multiple Sclerosis (RRMS) [25]. The procedures were approved by an Institutional Review Board and all participants who volunteered provided written informed consent. The sample was recruited through a research advertisement posted on the National MS Society (NMSS) website and distributed through 12 mid-western chapters of the NMSS. Those who were interested in the study contacted the research team by either e-mail or a toll-free telephone call. This contact was followed by a scripted conversation with the project coordinator, who described the study procedures and undertook screening for inclusion criteria. The inclusion criteria were: (1) diagnosis of RRMS confirmed by a physician; (2) relapse-free in the previous 30 days; (3) ambulatory with or without assistance (i.e., walk independently or walk with a cane or crutch or walker or rollator); and (4) willingness to complete the study materials every 6 months over 2.5 years. Those who did not satisfy the inclusion criteria were excluded from participation.

We successfully contacted 375 of the 463 people who expressed interest in the study, and 6 were uninterested in participation after the description of the study procedures. The remaining 369 people underwent screening, 44 did not satisfy the inclusion criteria, and 5 declined voluntary participation. We sent an informed consent document (completed by the participant) and RRMS verification form (completed by the participant's treating physician) to the remaining 320 people, and 41 did not return the documents despite 3 attempts for follow-up contact. We sent study materials to the remaining 279 people, and 10 subsequently declined further participation; this distribution of materials occurred in 12 waves of about 25 participants per wave beginning in March of 2008 (wave 1) and ending in February of 2009 (wave 12). There were 269 people with RRMS who provided baseline data. Of the initial 269 people, there were 258, 253, 245, 244, and 238 who provided follow-up data 6, 12, 18, 24, and 30 months later (i.e., 88%–96% of the initial sample). This attrition involved either a change in the participant's residential address or loss of materials through the US Postal Service.

2.2. Procedure

Participants were sent an accelerometer and battery of questionnaires through the U.S. Postal Service. We further provided pre-stamped and pre-addressed envelopes for return postal service. The project coordinator called to make sure the participants received the materials and understood the instructions. The participants then completed the battery of PRO questionnaires, and then wore the accelerometer for seven days. After completing the measures and wearing the accelerometer, participants returned the study materials through the U.S. Postal Service. We contacted participants by telephone and e-mail as a reminder to return the study materials up to 3 times. We further collected any missing questionnaire data based on follow-up telephone calls. This same procedure was completed every six-months over a 2.5-year period of time. All participants received $120 remuneration; this was prorated to be $20 per completion and return of the study materials.

2.3. Measures

2.3.1. PROs

For the purpose of this secondary analysis, we focused our attention on the responsiveness of the MSWS [16] and the MSIS [22]. The MSWS is a 12-item PRO measure of the impact of MS on walking. Scores range from 0- to 100, with higher scores reflecting greater impact of MS on walking. The MSIS is a 29-item PRO that assesses the physical (20 items) and psychological (9 items) impact of MS. Scores range from 0 to 100, with higher scores reflecting greater impact of MS on functioning.

Demographic data and the Patient-Determined Disease Steps (PDDS) [26] PRO were included to describe the sample. The PDDS is a self-report measure that was modeled after and correlates highly with the Extended Disability Status Scale [27]. This measure characterizes patient disability level into 1 of 9 steps (0, normal; 1, mild disability; 2, moderate disability; 3, gait disability; 4, early cane; 5, late cane; 6, bilateral support; 7, wheelchair or scooter; 8, bedridden) [26].

2.3.2. Performance-based measure

Community ambulation monitoring was done using the ActiGraph model 7164 accelerometer (ActivGraph, Pensacola, FL). This tool samples walking in the context of daily life where it naturally occurs to obtain ecologically 'valid' information (i.e., information that is generalizable to real-world, real-life experiences) [28]. This motion sensor is typically worn on a belt around the waist during the waking hours of everyday life over a period of 7 days. Recognized as a possible 'gold standard' measure of ambulation in MS [29], this device captures the overall ambulatory activity undertaken in one's usual environment and across the usual range of activities [29].

This brand of motion sensor further has acceptable accuracy across the disability spectrum (i.e., EDSS of 0–6.5 or bilateral device for ambulation) and a range of walking speeds (i.e., slow through comfortable and fast) in persons with MS [30]. The ActiGraph model 7164 accelerometer contains a single, vertical axis piezoelectric bender element that generates an electrical signal that is proportional to the force acting on it during ambulation. The acceleration/deceleration signal is digitized by an analog-to-digital converter and numerically integrated over a pre-programmed epoch interval. At the end of each interval,