

Evolution of Patient-Reported Outcomes and Their Role in Multiple Sclerosis Clinical Trials

Cindy J. Nowinski¹  · Deborah M. Miller² · David Cella¹

Published online: 14 September 2017

© The American Society for Experimental NeuroTherapeutics, Inc. 2017

Abstract Patient-reported outcomes (PROs) are playing an increasing role in multiple sclerosis (MS) research and practice, and are essential for understanding the effects that MS and MS treatments have on patients' lives. PROs are captured directly from patients and include symptoms, function, health status, and health-related quality of life. In this article, we review different categories (e.g., generic, targeted, preference-based) of PRO measures and considerations in selecting a measure. The PROs included in MS clinical research have evolved over time, as have the measures used to assess them. We describe findings from recent MS clinical trials that included PROs when evaluating Food and Drug Administration-approved disease-modifying therapies (e.g., daclizumab, teriflunomide). Variation in the measures used in these trials makes it difficult to draw any conclusions from the data. We therefore suggest a standardized approach to PRO assessment in MS research and describe 2 generic, National Institutes of Health-supported measurement systems [Neuro-QoL and the Patient-Reported Outcomes Measurement Information System (PROMIS)] that would facilitate such an approach. The use of PROs in MS care and research is expanding beyond clinical trials, as is demonstrated by examples from comparative effectiveness and other patient-centered research. The importance of PRO assessment is expected to continue to grow in the future.

Keywords Patient-reported outcomes · Health-related quality of life · Multiple sclerosis · Clinical trials

Patient-reported outcomes (PROs) refer to direct patient reports of their experience with a disease and its treatment. PROs are a subset of the broader concept of “patient-generated outcomes” that refer to the full range of health information provided by patients about themselves, which can include demographic information, prior history, and so on. PROs are obtained via “report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else” [1]. PROs include symptoms, functional level, health status, health utility, and health-related quality of life (HRQoL). By their very nature, PROs are patient-centered and their addition to traditionally collected anatomical, biological, and clinical data has resulted in a fundamental shift in how research and clinical practice are conducted. This article presents an overview of PRO assessment, how its inclusion in multiple sclerosis (MS) clinical research has evolved over time, and directions for the future.

MS is a chronic progressive disease that usually manifests in young adulthood. It has limited impact on life expectancy. There is no typical course to disease progression and MS produces a myriad of symptoms that can occur at any time and in different combinations. MS is not a curable disease; treatment tends to focus on symptom management, limiting disease progression with the use of disease-modifying drugs (DMDs), and maximizing quality of life. MS research, then, must include evaluation of how well interventions achieve these goals. As patients appreciate the totality of their disease experience, these goals can best be defined with their input. In addition, patient perceptions frequently differ from those of clinicians [2–5]. Assessment of patient functioning in the

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13311-017-0571-6>) contains supplementary material, which is available to authorized users.

✉ Cindy J. Nowinski
c-nowinski@northwestern.edu

¹ Northwestern University Feinberg School of Medicine, Chicago, IL, USA

² Mellen Center, Cleveland Clinic, Cleveland, OH, USA

treatment setting, however well measured, may not accurately reflect their functioning at home [6]. Thus, PROs are generally accepted as important to assess in clinical research for most chronic conditions [7, 8], including MS (see the proceedings from the annual meetings of the International Society for Quality of Life Research) [9–11]. Further, PROs are increasingly seen as supporting every aspect of the healthcare continuum from research to clinical practice (including clinical decision-making and quality reporting) through to public health [12–16].

Approaches to PRO Assessment

Most PRO measures are categorized as either generic or targeted. *Generic* measures include questions that are general enough for use with both healthy and clinical populations. Generic measures used in MS include the Medical Outcomes Study Short Form-36 (SF-36) [17], the Sickness Impact Profile [18], and versions of the Health Utilities Index [19]. *Targeted* measures are comprised of questions aimed towards specific diseases (e.g., MS), domains (e.g., cognition, fatigue), or interventions (e.g. use of biological response modifiers). Examples of symptom-focused, domain-specific measures include the Brief Pain Inventory [20], and the pain, fatigue, depression, sleep, and other symptom measures included in the Patient-Reported Outcomes Measurement Information System (PROMIS) and the Neurology Quality of Life (Neuro-QoL) measurement system. Disease-specific measures for MS are numerous, and include the Multiple Sclerosis International Quality of Life (MusiQoL) [21], the Multiple Sclerosis Quality of Life-54 (MSQOL-54) [22], the Functional Assessment of Multiple Sclerosis (FAMS) [23], the Multiple Sclerosis Impact Scale (MSIS-29) [24], the Patient Reported Impact of Multiple Sclerosis (PRIMUS) [25], the Hamburg Quality of Life Questionnaire in Multiple Sclerosis [26], the MS Quality of Life Inventory [27], the Multiple Sclerosis Impact Profile [28], the Leeds Multiple Sclerosis Quality of Life scale [29], the Disability and Impact Profile [30], and the RAYS scale [31]. Some PRO measures, like the Disability and Impact Profile, also incorporate patients' perceptions of how important each effect of MS is on their lives. Generic PROs, usually normed against general/healthy populations, are appropriate for cross-disease comparisons, and are useful in resource allocation and cost-effectiveness analyses. In contrast, targeted measures can provide more in-depth and comprehensive coverage of a specific domain or area, are thought to be more sensitive to changes in health status or function, and questions may seem more relevant and therefore more acceptable to patients [32]. This has led to recommendations for using a combined approach (generic and targeted) when feasible [33, 34]. The MSQOL-54, for example, includes the generic SF-36 along with 18 MS-specific questions. PROs also differ in whether they measure a single dimension of health or multiple dimensions.

Multidimensional, or profile measures, typically provide separate scores for each dimension rather than a summary score. Some, such as the FAMS, provide both individual dimension scores and a total summary score.

Health Utility Measures

Health utility measures are a special type of PRO derived from economic and decision theory. Utility measures cover multiple health domains but provide a single summary score. This score reflects patient preferences for different health states, with higher preferences signifying greater value or desirability. Scores typically range from 0 (death) to 1 (perfect health), although negative numbers (states worse than death) are possible. Utility scores can be used to compute quality-adjusted survival, measured in quality-adjusted life years (QALYs), by supplying the “quality” component of the equation. QALYs combine both mortality (survival) and morbidity (HRQoL) into a single index [35, 36], applicable across diseases and interventions. QALYs are frequently used in cost-effectiveness and similar analyses to help guide resource allocation. Commonly used generic utility measures include the EQ-5D [37], the Quality of Well-Being scale [38], and the Health Utility Index [39, 40]. However, generic health utility measures have been criticized for being insensitive to treatment effects and/or failing to adequately cover all relevant dimensions for a given condition [41]. This has led to efforts to develop disease- or condition-specific utility measures such as the 15D [42] and the Multiple Sclerosis Impact Scale – Eight Dimensions [43, 44].

Item Response Theory and PRO Assessment

In recent years, PROs have begun to incorporate modern measurement science [e.g., item response theory (IRT)] into their development. Utilization of IRT provides the instruments with certain advantages, including the ability to be brief while remaining precise and valid [45]. Using IRT methodology, sets of questions (items) are calibrated along a continuum that covers the full range of the construct to be measured. Once calibrated, any or all items in this “bank” can be used to generate a score. Users can select specific items to create “short forms” (SFs), typically consisting of 6 to 8 questions that meet their measurement needs. For example, a user wishing to assess upper extremity function in a group of patients who tended to have poor or very poor fine motor abilities could create a custom SF primarily comprised of those items targeting lower levels of upper motor function that cluster near the lower end of that bank. Item banks are also the basis for computerized adaptive testing (CAT). This is a specialized type of computer-based testing in which, after the initial item is presented, the test administration algorithm selects each item to be presented based on the response to the previous

item. Items are adaptively selected until either the desired level of score precision is achieved (typically after 4–6 items) or a predetermined maximum number of items (e.g., 12) have been administered. CAT allows for frequent and precise evaluation of patients at the individual level while placing minimal burden on patients [45–48]. Users can administer short, unique tests to every individual, with reliability and scores equivalent to longer, fixed-length assessments.

The Need for Common PRO Measures: Neuro-QoL and PROMIS

The variety of generic and MS-specific measures available for use in MS research has some benefits, but the lack of a set of standard measures also has significant disadvantages. For example, failure to use common measures results in an inability to combine and compare data across different studies, which slows the pace of discovery. Further, some available measures are of uncertain validity and were created without using modern test development methodology. This state of affairs led the National Institutes of Health (NIH), in the mid-2000s, to support the creation of 2 standard sets of PRO measures, 1 appropriate for use across neurological conditions (Neuro-QoL) and 1 for use across a broad range of chronic health conditions (PROMIS). These systems are described in more detail below.

Neuro-QoL (Quality of Life in Neurological Disorders)

Neuro-QoL is a clinically relevant, validated PRO measurement system to assess the HRQoL of adults and children with neurological conditions [49–51]. Neuro-QoL development was funded by the National Institute of Neurological Disorders and Stroke (NINDS) with the goal of creating a standard set of measures to assess outcomes relevant across common neurological conditions (generic), as well as outcomes relevant to specific patient populations (targeted) that would permit comparisons across neurological populations.

The development methodology used for Neuro-QoL, which included incorporation of patient input and utilization of IRT, is consistent with Food and Drug Administration guidance regarding PROs and the European Medicines Agency Reflection Paper on the use of HRQL measures [52]. The adult Neuro-QoL consists of 12 item banks and 1 scale assessing important aspects of mental well-being (anxiety; depression; positive affect and well-being; cognitive function; emotional and behavioral dyscontrol; communication); physical well-being (upper extremity function – fine motor, activities of daily living; lower extremity function - mobility; sleep disturbance; fatigue; stigma), and social well-being (ability to participate in social roles and activities; satisfaction with social roles and

activities). These areas were chosen using input from patients, caregivers, and clinical providers [53]. The item banks are the basis for standalone fixed-length SFs and CAT. The standard Neuro-QoL SFs, typically consisting of 8 to 9 items selected with clinical expert input, were validated in 5 major neurological adult conditions, including MS. With respect to MS, the SFs demonstrated good internal consistency, test–retest reliability, and concurrent and known groups validity [54]. Some initial evidence of responsiveness to self-reported change was shown in the initial Neuro-QoL observational trial. Ongoing efforts (e.g., in the MS PATHS initiative, described later in this paper) to collect longitudinal Neuro-QoL data along with other disease characteristics will enable closer evaluation of responsiveness. Neuro-QoL data are reported as T-scores (mean of 50 and SD of 10) referenced to either general or clinical population samples. T-scores enable users to determine how far a given score is from “average”. For some measures, additional interpretative information regarding minimal values reflecting important change and cut points for different severity levels is available [55].

PROMIS

What began as an NIH Roadmap initiative to create a publicly available system for measuring generic PROs across diseases and conditions, PROMIS currently offers an extensive library of measures available in a range of flexible assessment options [56]. PROMIS item banks, with associated CATs and short forms, cover common symptoms, functions, behaviors, and feelings within the overarching domains of physical, mental, and social well-being. While some Neuro-QoL and PROMIS measures cover the same domains (in which case the Neuro-QoL measure is preferred since it was validated in neurological conditions), PROMIS dramatically expands the range of PROs that can be assessed. As with Neuro-QoL, PROMIS measures are IRT-based, were developed in a manner consistent with regulatory agencies’ PRO guidance, and have been shown to be valid, reliable, and responsive [57, 58]. Recently, a process to derive a generic preference-based summary score (PROMIS-Preference or PROPr) from PROMIS measures has been developed (<http://janelhanmer.pitt.edu/ProPr.html>) and is available for use.

Both Neuro-QoL and PROMIS measures convey several advantages. They are brief (~1.5 min per short form; ≤ 1 min per CAT [59]), available in a variety of administration formats (interviewer or self; paper-and-pencil, computer), and patient centered (e.g., assess outcomes that patients identified as important). When using CATs, investigators can achieve precise scoring at the individual patient level while maintaining brevity. Distributed through HealthMeasures (www.healthmeasures.net), a resource for disseminating and supporting 4 NIH-supported measurement systems, PDFs of all Neuro-QoL and

PROMIS measures can be freely downloaded from the HealthMeasures website. They are also contained within data-collection tools that provide automated scoring and immediate data access [60]. All measures are available in English and Spanish, with additional translations available for individual measures (e.g., > 15 for fatigue). Finally, many Neuro-QoL and PROMIS measures are “linked” to each other through efforts of the PROsetta Stone (www.prosetta.org) project. Linking provides equivalent scores for different measures of the same health outcome through an IRT-based mechanism of equating the instruments along a common measurement continuum [61]. Using tables provided by PROsetta Stone, users can convert scores on Neuro-QoL to scores on PROMIS and vice versa.

Neuro-QoL and PROMIS also have some limitations. They do not cover all domains relevant to MS, only some of the measures have been validated in MS, and the availability of translations varies across instruments. Further, additional work is needed to demonstrate sensitivity to change and the amount of change that is clinically significant.

Selecting a PRO

In selecting a PRO measure, the first step is to identify candidate PROs that have a conceptually valid link to the outcome(s) of interest. As noted above, however, this may result in a large number of PROs from which to choose. Fortunately, guidance on selecting PROs is available [62–64]. In general, it is important to consider the setting (e.g., clinic), the purpose for which the PRO will be used, and characteristics of the instrument itself. One should evaluate both the psychometric evidence for using a measure, as well as consider practical aspects such as ease of use (e.g., respondent burden, data collection options) and other logistics related to implementation [65]. The degree to which a measure conforms to regulatory guidelines for measure development, such as patient involvement in domain selection and question refinement, should be assessed. Francis and colleagues (2016) recently synthesized available guidelines and recommendations into a checklist that clinicians and researchers at varying levels of expertise can use to evaluate how well a PRO meets widely accepted criteria for selection: conceptual model, content validity, reliability, construct validity, scoring and interpretation, and respondent burden [66]. Cultural relevance, linguistic adaptation, and respondent literacy level may be particularly important in multinational or multi-regional trials [10].

A NINDS initiative, the Common Data Elements (CDE) project (<http://www.commondataelements.ninds.nih.gov>), can also be helpful in measure selection. The overarching, ongoing goal of the CDE project is to standardize data collection in neurology clinical research, including collection

of PROs such as ability to perform activities of daily living and HRQoL. NINDS CDEs for MS were released in 2011 and are periodically updated.

Use of PROs in MS Clinical Research

Assessment of PROs is of critical importance in research on the consequences of MS, as well as treatments for the disease. Typically, more effective DMDs also tend to carry a greater risk of treatment adverse events [67]. As the evaluation of PROs has increased in clinical studies, our understanding of how MS and MS treatments affect patients has also increased. In their seminal 2003 review article on the impact of MS on quality of life, Benito-Leon et al. [68] described the first published paper, which appeared in 1992 [69], that assessed how living with MS affects HRQoL. It demonstrated that individuals living with MS had more impaired HRQoL than those living with inflammatory bowel disease or rheumatoid arthritis. They noted that through the 1990s, research on the impact of MS was primarily limited to assessment of impairment and disability especially as measured by the Expanded Disability Status Scale [70, 71]. In the decades that followed, a variety of PROs were incorporated into MS research and measured in varying ways. PROs have been used to demonstrate the negative impact of MS *versus* the general population (using the generic measures WHO-QOL) [72–74], and to show how PROs evolve over the course of MS through registries [75–78] (using the SF-12 [79] and the PROMIS-10) and in observational studies (using the SF-36, the Leeds Multiple Sclerosis Quality of Life, and the MusiQOL) [80–82]. PROs also provide unique insight into the negative effect MS symptoms have on aspects of well-being, including pain (assessed using the SF-36, the MSQOL-54, and the FAMS) [83–85]; cognition (using the SF-36, the FAMS and the MusiQOL) [86–88]; sexual dysfunction (using the SF-36 and the MSQOL-54) [89–91], and depression (using the SF-36 and the MSQOL-54) [92, 93].

PROs as Secondary Endpoints in Randomized Clinical Trials of DMDs

As reported by the Multiple Sclerosis Coalition, as of March 2017, there are currently 14 disease-modifying agents that have been approved by the appropriate regulatory agencies. Of these 14, the PRO results from randomized controlled phase III trials for 3 of the medications reported before 2010 have been previously summarized [94, 95]. The PROs used in those studies included both generic and disease specific measures such as the generic Sickness Impact Profile [interferon (IFN)- β 1b [96]] and SF-36 (natalizumab [97]) as well as the

disease specific FAMS (IFN- β 1b [98]) and the MS Quality of Life Inventory (IFN- β 1b [99] and IFN- β 1a [100]).

In order to update the PRO research in phase III clinical trials, the authors of this review, in March 2017, conducted an updated literature search of PROs used in randomized phase III trials of MS DMDs. It was based on searches of MEDLINE and Pubmed, from January 2011 to March 2017, using the search terms “multiple sclerosis”, “randomized trial”, “disease modifying therapy”, “quality of life”, “patient reported outcomes”, and “PROs”. The recent review of DMDs and PROs by Jongen [101] also aided in this search. These searches identified 5 phase III, randomized controlled trials of MS DMDs with PRO endpoints.

Two injectable medications have recently been approved using data from these trials: pegylated (peg) IFN- α [102, 103] and daclizumab [104, 105]. pegIFN- β 1a 125 μ g is administered subcutaneously every 2 weeks and has been proved an effective treatment *versus* placebo or to pegIFN- β 1a 125 μ g administered every 4 weeks. The PRO benefit of pegIFN was demonstrated in the ADVANCE study by comparing the MSIS-29 scores at 42 weeks with MSIS-29 baseline scores in the 3 treatment groups: placebo-treated patients, those treated with pegIFN- β 1a 125 μ g administered every 2 weeks and pegIFN- β 1a 125 μ g administered every 4 weeks [106]. In the placebo group, based on the MSIS-29 physical subscale, there was a statistically significant worsening from baseline to month 48 [mean 1.24, 95% confidence interval (CI) 0.05–2.44]. In the pegIFN- β 1a every 2 and every 4 weeks groups, no statistically significant worsening was found on the MSIS-29 physical subscale, which had mean changes of 0.08 (95% CI –1.10 to 1.27) and 1.12 (95% CI –0.05 to 2.28), respectively. All 3 treatment groups were found to have a statistically significant improvement at week 48 from baseline on the MSIS-29 psychological subscale, with a mean change of –2.17 (95% CI –3.63 to –0.70), –2.06 (95% CI –3.58 to –0.53), and –1.70 (95% CI –3.24 to –0.15) in the placebo and pegIFN- β 1a every 2 and every 4 weeks groups, respectively. The between-group differences in mean change from baseline were not statistically significant ($p > 0.05$) for MSIS-29 physical or psychological subscales, and were smaller than differences considered to be clinically meaningful.

Daclizumab, the other recently approved self-injected DMD, is administered at a dose of 150 mg once monthly. The SELECT trial was designed to assess if daclizumab high-yield process (HYP) at 2 doses (150 mg and 300 mg) was effective when given as a monotherapy [104]. The MSIS-29, the SF-12, and the EQ-5D were included as PRO tertiary endpoints. The 2 doses of daclizumab HYP were superior to placebo on the primary and the majority of the additional endpoints. For the PROs, there was significant improvement in the mean MSIS-physical subscale in the group receiving the 150 mg (mean \pm SD change score -1.0 ± 11.8) compared with placebo (mean \pm SD 3.0 ± 13.5 ; $p = 0.00082$), whereas there

was no statistical difference between 300 mg and placebo ($p = 0.13$). Similar patterns of significance were found for the MSIS-psychological subscale and the SF-12 physical and mental components. Only with the EQ-5D visual analog scale was there a significant difference between the 300 mg group and placebo ($p = 0.015$), although not as significant as the difference between 150 mg and placebo ($p < 0.001$). A post-hoc multivariate analysis of these data demonstrated that relapse occurrence and confirmed disability progression at week 12 were statistically associated with decline in PRO scores [107]. In the second trial of daclizumab HYP, it was compared with IFN- β 1a in relapsing MS. Participants in this DECIDE trial had greater disease activity than those in the SELECT trial [108]. Patients were randomized at a 1:1 ratio to receive 150 mg subcutaneous daclizumab and intramuscular placebo monthly or 30 μ g intramuscular IFN- β 1a and subcutaneous placebo monthly. The primary endpoint was annualized relapse rate. The MSIS-29 was included as a tertiary endpoint. The study was positive for daclizumab on the primary endpoint and MRI results, but there was no between-group difference on the MSIS-29.

Three oral MS DMDs, teriflunomide [109, 110], fingolimod [111], and dimethyl fumarate [105, 112] have recently been approved and the pivotal trials include PRO data. During the initial teriflunomide phase III trial (TEMPO), eligible patients were randomly assigned to receive a once-daily oral dose of placebo, 7 mg teriflunomide, or 14 mg teriflunomide for 108 weeks [110]. The primary endpoint was reduction in annualized relapse rates. The PRO Fatigue Impact Scale (FIS) was included as a secondary endpoint [113]. The primary endpoint was met for both doses of teriflunomide. However, patients in all groups reported very limited change in fatigue from baseline and there were no significant between-group differences in those changes [teriflunomide 7 mg *vs* placebo ($p = 0.39$); teriflunomide 14 mg *vs* placebo ($p = 0.83$)]. The second teriflunomide trial, TOWER, was designed to provide additional safety and efficacy data [109]. Study duration was a fixed time point 48 weeks after the last patient was randomized. The trial had similar inclusion criteria, treatment assignments, and primary and secondary endpoints as the TEMPO study. In addition to the FIS, the SF-36 was included as a PRO in this trial. These PRO data were collected at week 48 and last study visit. In the primary analysis, teriflunomide 7 mg and teriflunomide 14 mg significantly reduced the annualized relapse rate compared with placebo. For both Physical Component Score (PCS) and Mental Component Summary (MCS) of the SF-36, neither treatment differed from placebo during the 48-week treatment. There was no between-group difference from baseline to last visit on the SF-36 PCS score, whereas there was a significant between-group difference favoring teriflunomide 14 mg to placebo during the time frame ($p = 0.02$) for the SF-36 MCS score. There were no between-group differences

on FIS score at week 48, whereas there was a significant between-group FIS difference favoring teriflunomide 14 mg to placebo from baseline to last study visit ($p = 0.04$).

The FREEDOMS II trial [111] was the third pivotal trial for fingolimod following the TRANSFORMS [114] and FREEDOMS I [114] trials. It was designed to further assess the drug's safety and efficacy at doses of 0.5 mg and 1.25 mg *versus* placebo once daily. The primary endpoint was annualized relapse rate after 24 months of treatment. The PRO secondary endpoints included the EQ-5D, the PRIMUS, and the mFIS. While the primary endpoint of annualized relapse rate was reduced in the treatment groups compared to placebo ($n = 778$), there were no significant between-group differences for any of the PRO scores.

The effects of dimethyl fumarate on PROs in CONFIRM, a randomized placebo-controlled study, were evaluated in a secondary analysis of that trial [112]. The investigators used the SF-36 to assess the between-group differences between dimethyl fumarate 240 mg twice daily, dimethyl fumarate 240 mg 3 times daily, and glatiramer acetate 20 mg 4 times daily. Change in SF-36 scores over 2 years was the target PRO endpoint. At study's end, a higher proportion of all subjects in the active treatment groups had a clinically significant change (≥ 5 point) in both SF-36 PCS and MCS. This change was also statistically significant for the dimethyl fumarate 3 times daily group.

The CARES-MS I and CARES-MS II studies compared alemtuzumab to IFN- β 1a in unblinded phase III trials treating relapsing-remitting MS. These trials included FAMS as the primary PRO and also included the SF-36 and EQ-5D. Results indicated significant and sustainable advantages to alemtuzumab on all 3 measures [115]. In both studies, patients on alemtuzumab significantly improved from baseline to month 6 on 5 of 6 FAMS components (mobility, symptoms, thinking and fatigue, general contentment, and emotional well-being), and improvements were maintained until month 24 ($p < 0.05$). Between-group comparisons of change in FAMS component scores favored alemtuzumab over subcutaneous IFN- β 1a at multiple time points [mobility, symptoms, and thinking and fatigue (both studies), and general contentment (CARE-MS II only)]. Our search found no reports of PROs included as outcome measures in studies of the following injectable agents: IFN- β 1a subcutaneous (Rebif), generic IFN- β 1b subcutaneous (Extavia), glatiramer acetate or its generic equivalent (Glatopa), or the infusion drug mitoxantrone.

Based on the variability of PRO data from these trials and the variety of measures implemented it is not possible to reach any conclusion about the overall benefit of DMDs or the relative benefit of one over another on the well-being of individuals living with MS. As described earlier in this paper, the availability of standardized measures, including PROMIS and Neuro-QoL, that assess general well-being as well as

discrete aspects (e.g., symptoms, function) of physical, psychological, and social well-being offers a systematic approach to assessing patient well-being in clinical trials and other research methodologies. Use of a standard measurement approach will allow for PRO comparisons across DMD trials and the determination of which DMDs are most appropriate for differing cohorts of patients with MS, i.e. for the purposes of comparative effectiveness research (CER), which is discussed in more detail below.

As is suggested in the above review, the primary endpoints in the majority of clinical trials are of 2 types. The first are clinical endpoints such as annualized relapse rate or time to confirmed disease progression. The second type are biomarker endpoints that measure the level of ongoing disease activity. However, as demonstrated in the above review, there has been growing recognition of the importance of patient-reported endpoints.

Expanding Research Approaches

Many countries use health technology assessments (HTA) to systematically evaluate the properties, impacts, and benefits/added value of health technologies, including medications [116, 117]. HTA results are used to inform healthcare policy and decision-making, such as reimbursement decisions and the establishment of clinical guidelines. HTA agencies often emphasize the involvement of patient and other stakeholder perspectives in the HTA process, including measurement of and consideration of HRQoL as part of a technology evaluation [118]. A related research approach that is gaining increasing acceptance in the USA, especially for chronic illnesses, is CER. As defined by the Patient-Centered Outcomes Research Institute (PCORI), which was created by the US legislation known as the Patient Protection and Affordable Care Act (ACA), CER includes research approaches that help patients and other healthcare stakeholders, including caregivers, clinicians, insurers, and others, make better-informed decisions about their health and healthcare [119]. CER methodologies include randomized clinical trials, large pragmatic studies, and large-scale observational studies. In 2015, PCORI convened a MS Stakeholder Workgroup that established the 5 top CER priorities [120]. These included: 1) What are the comparative benefits and harms of nonpharmacological and pharmacological approaches in relation to key symptoms (e.g., emotional health, fatigue, cognition, pain) in people with MS? 2) In people with progressive MS, what is the comparative effectiveness of different care delivery approaches (i.e., MS specialty center *vs* community neurology; direct care *vs* telemedicine; "specialized medical home" *vs* community neurology delivery of care) in improving outcomes such as functional status, quality of life, symptoms, emergency room use, and hospitalization? 3) Does an integrative model of care along

with DMDs in newly diagnosed individuals affect disability progression and symptoms (physical, emotional, and cognitive) compared with treatment with DMD alone? 4) Among patients with MS receiving a DMD who experience disease activity, what are the benefits and harms of continuing the same therapy *versus* changing to a new medication? 5) What are the comparative benefits and harms of different disease-modifying therapies in newly diagnosed relapsing-remitting MS on disease activity, disease progression, symptoms, and quality of life?

Members of the PCORI MS Stakeholders Workgroup and others recommend the consideration of additional types of study endpoints, such as composites, in order to expand the concepts of interest beyond the commonly used clinical endpoints named above. One such composite, comprised of 3 dimensions of MS disability, is the Multiple Sclerosis Functional Composite, which assesses ambulation, manual dexterity, and cognition—all clinical outcome measures [121–124]. Another such composite, “no evidence of disease activity” (NEDA) is being put forth as the goal of MS treatment [11]. It currently consists of Expanded Disability Scale Score disease progression, relapse rate, and formation of MRI lesions, and may include any measure of disease activity. Many argue that this definition of NEDA is far too narrow and needs to include other parameters, including PROs [11, 125–127].

If the objective of MS care is NEDA, and it is to be demonstrated in a meaningful time frame, that outcome can most effectively be demonstrated by the large-scale observational studies proposed as one of the CER methodologies by PCORI. An effective way to develop the infrastructure for conducting methodologically rigorous observational studies is through the creation and maintenance of a “learning healthcare system” as is described in the Institute of Medicine’s “Best Care at Lower Cost. The path to continuously learning health care in America” [128]. The Institute of Medicine provides the following definition:

A learning health care system is one in which science, informatics, incentives, and culture are aligned for continuous improvement and innovation, with best practices seamlessly embedded in the care process, patients and families active participants in all elements, and new knowledge captured as an integral by-product of the care experience (Roundtable on Value & Science-Driven Health Care, 2012).

In order to support the CER approach, the Cleveland Clinic Mellen Center for Multiple Sclerosis Treatment and Research has partnered with 8 other MS comprehensive care centers to form MS PATHS, the first learning health system for MS [129]. The aims of MS PATHS is to create a network of healthcare institutions, leveraging technology and patient

engagement to collect standardized, quantitative data on each patient attending the member institution at every follow-up clinical encounter. Willing patients who agree to share anonymized data are enrolled in the data collection process. This project is anticipated to use quantitative, multidimensional data to advance disease understanding, support effectiveness research, accelerate translational research, and provide outcomes data for value-based models of reimbursement. The computerized adaptive testing method of the Neuro-QoL assessment platform is the instrument used to obtain PRO data across the MS PATHS sites.

Conclusion

Implementing PROs has become an important consideration for a number of agencies involved in generating basic and clinical research, financing healthcare, and vetting endpoints that lead to drug and device approval. In the USA, the NIH and the NINDS have devoted considerable resources to the development and validation of PROs and intend their use as study endpoints in human subject trials they support [130]. PROs are an important component of the Affordable Care Act. Under that act, the Centers for Medicare and Medicaid have included PROs in their Quality Development Plan: Supporting the Transition to the Merit-based Incentive Payment System and the Alternative Payment Models [131] and its mandate to demonstrate quality in order to enhance reimbursement (pay for performance) [132].

Clearly, PROs are core outcomes of MS therapy today. Improvements in HRQoL and other PROs are possible, but not always obtained, with today’s MS treatments. There are many available measurement options, including generic, symptom-targeted, and MS-specific, available to researchers and clinicians for use when assessing these outcomes. Available tools now include modern measures that utilize the assets of item response theory, including custom short forms or computerized adaptive testing, to tailor assessment of life domains that are important to people living with MS. Going forward, clinical trials and comparative effectiveness research can help determine which of the many available therapies approved for treatment of MS is likely to be most effective for people on the outcomes they care most about.

Acknowledgments This work was supported by funding from the National Cancer Institute, (Grant# 1U2CCA186878-01) and previously by the National Institute of Neurological Disorders and Stroke (Contract #HHSN271201200036C).

Required Author Forms Disclosure forms provided by the authors are available with the online version of this article.

References

1. US Food and Drug Administration. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. Available at: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM071975.pdf>. Accessed 15 June 2017.
2. Edelman EJ, Gordon K, Justice AC. Patient and provider-reported symptoms in the post-cart era. *AIDS Behav* 2011;15:853-861.
3. Janse AJ, Gemke R, Uiterwaal C, et al. Quality of life: patients and doctors don't always agree: a meta-analysis. *J Clin Epidemiol* 2004;57:653-661.
4. Kremenutzky M, Walt L. Perceptions of health status in multiple sclerosis patients and their doctors. *Can J Neurol Sci* 2013;40:210-218.
5. Rothwell PM, McDowell Z, Wong CK, Dorman PJ. Doctors and patients don't agree: Cross sectional study of patients' and doctors' perceptions and assessments of disability in multiple sclerosis. *BMJ* 1997;314:1580.
6. Freeman JA, Langdon DW, Hobart JC, Thompson AJ. Inpatient rehabilitation in multiple sclerosis do the benefits carry over into the community? *Neurology*.1999;52:50.
7. Doward LC, Gnanasakthy A, Baker MG. Patient reported outcomes: looking beyond the label claim. *Health Qual Life Outcomes* 2010;8:89.
8. Sonnad SS, Mullins CD, Whicher D, et al. Recommendations for the design of phase 3 pharmaceutical trials that are more informative for patients, clinicians, and payers. *Contemp Clin Trials* 2013;36:356-361.
9. Hareendran A, Gnanasakthy A, Winnette R, Revicki D. Capturing patients' perspectives of treatment in clinical trials/drug development. *Contemp Clin Trials* 2012;33:23-28.
10. Gnanasakthy A, Lewis S, Clark M, Mordin M, DeMuro C. Potential of patient-reported outcomes as nonprimary endpoints in clinical trials. *Health Qual Life Outcomes* 2013;11:83.
11. van Munster CE, Uitdehaag BM. Outcome measures in clinical trials for multiple sclerosis. *CNS Drugs* 2017;31:217-236.
12. Ahmed S, Berzon RA, Revicki DA, et al. The use of patient-reported outcomes (PRO) within comparative effectiveness research: Implications for clinical practice and health care policy. *Med Care* 2012;50:1060-1070.
13. Baumstarck K, Boyer L, Boucekine M, et al. Measuring the quality of life in patients with multiple sclerosis in clinical practice: a necessary challenge. *Mult Scler Int* 2013;2013: 524894.
14. Coulter A. Measuring what matters to patients. *BMJ* 2017;356:j816.
15. Miller DM, Allen R. Quality of life in multiple sclerosis: determinants, measurement, and use in clinical practice. *Curr Neurol Neurosci Rep* 2010;10:397-406.
16. Van der Wees PJ, Nijhuis-Van Der Sanden MW, Ayanian JZ, et al. Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. *Milbank Q* 2014;754-775.
17. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-483.
18. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The sickness impact profile: Development and final revision of a health status measure. *Med Care* 1981;19:787-805.
19. Feeny DH, Torrance GW, Furlong WJ, Spilker B. *Health utilities index*. Philadelphia, PA: Lippincott-Raven; 1996.
20. Cleeland C. Measurement of pain by subjective report. *Adv Pain Res Ther* 1989;12:391-403.
21. Simeoni MC, Auquier P, Fernandez O, et al. Validation of the multiple sclerosis international quality of life questionnaire. *Mult Scler* 2008;14:219-230.
22. Vickrey BG, Hays RD, Harooni R, Myers LW, Ellison GW. A health-related quality of life measure for multiple sclerosis. *Qual Life Res* 1995;4:187-206.
23. Cella DF, Dineen K, Arason B, et al. Validation of the functional assessment of multiple sclerosis quality of life instrument. *Neurology* 1996;47:129-139.
24. 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;124:962-973.
25. Doward LC, McKenna SP, Meads DM, Twiss J, Eckert BJ. The development of patient-reported outcome indices for multiple sclerosis (PRIMUS). *Mult Scler* 2009;15:1092-102.
26. Gold SM, Heesen C, Schulz H, et al. Disease specific quality of life instruments in multiple sclerosis: Validation of the hamburg quality of life questionnaire in multiple sclerosis (HAQUAMS). *Mult Scler* 2001;7:119-130.
27. Fischer JS, LaRocca NG, Miller DM, et al. Recent developments in the assessment of quality of life in multiple sclerosis (MS). *Mult Scler* 1999;5:251-259.
28. Wynia K, Middel B, Van Dijk JP, et al. The multiple sclerosis impact profile (MSIP). Development and testing psychometric properties of an icf-based health measure. *Disabil Rehabil* 2008;30:261-274.
29. Ford HL, Gerry E, Tennant A, et al. Developing a disease-specific quality of life measure for people with multiple sclerosis. *Clin Rehabil* 2001;15:247-258.
30. Lankhorst GJ, Jelles F, Smits RC, et al. Quality of life in multiple sclerosis: the disability and impact profile (DIP). *J Neurol* 1996;243:469-474.
31. Rotstein Z, Barak Y, Noy S, Achiron A. Quality of life in multiple sclerosis: development and validation of the "rays" scale and comparison with the SF-36. *Int J Qual Health Care* 2000;12:511-517.
32. Özakbas S, Akdede BB, Kösehasanogullari G, Aksan Ö, Idiman E. Difference between generic and multiple sclerosis-specific quality of life instruments regarding the assessment of treatment efficacy. *J Neurol Sci* 2007;256:30-34.
33. Guyatt GH, Jaeschke R, Feeny DH, Patrick DL. Measurements in clinical trials: Choosing the right approach. In: Spilker B, editor. *Quality of life and pharmacoeconomics in clinical trials*. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 1996. p. 41-49.
34. Cella DF, Wiklund I, Shumaker SA, Aaronson NK. Integrating health-related quality of life into cross-national clinical trials. *Qual Life Res* 1993;2:433-440.
35. Hanmer J, Feeny D, Fischhoff B, et al. The PROMIS of QALYS. *Health Qual Life Outcomes* 2015;13:122.
36. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. *Br Med Bull* 2010;96:5-21.
37. EuroQol Research Foundation. EQ-5d. Available at: <https://euroqol.org/>. Accessed 19 June 2017
38. Kaplan RM, Anderson JP, Ganiats TG. The quality of well-being scale: Rationale for a single quality of life index. In: Walker SR, Rosser RM, editors. *Quality of life assessment: Key issues in the 1990s*. Amsterdam: Springer; 1993. p. 65-94.
39. Furlong WJ, Feeny DH, Torrance GW, Barr RD. The health utilities index (HUI®) system for assessing health-related quality of life in clinical studies. *Ann Med* 2001;33:375-384.
40. Horsman J, Furlong W, Feeny D, Torrance G. The health utilities index (HUI®): concepts, measurement properties and applications. *Health Qual Life Outcomes* 2003;1:54.
41. Revicki DA, Lenderking WR. Methods and issues associated with the use of quality-adjusted life-years. *Expert Rev Pharmacoecon Outcomes Res* 2012;12:105-114.

42. Dagklis IE, Aletras VH, Tsantaki E, Orologas A, Niakas D. Multiple sclerosis patients valuing their own health status: valuation and psychometric properties of the 15D. *Neurol Int* 2016;8: 6416.
43. Goodwin E, Green C. A quality-adjusted life-year measure for multiple sclerosis: developing a patient-reported health state classification system for a multiple sclerosis-specific preference-based measure. *Value Health* 2015;18:1016-1024.
44. Goodwin E, Green C, Spencer A. Estimating a preference-based index for an eight-dimensional health state classification system for multiple sclerosis. *Value Health* 2015;18:1025-1036.
45. Chang CH, Reeve BB. Item response theory and its applications to patient-reported outcomes measurement. *Eval Health Prof* 2005;28:264-282.
46. Revicki DA, Cella DF. Health status assessment for the twenty-first century: Item response theory, item banking and computer adaptive testing. *Qual Life Res* 1997;6:595-600.
47. Cella D, Chang CH. A discussion of item response theory (IRT) and its applications in health status assessment. *Med Care* 2000;38:1166-1172.
48. Hahn EA, Cella D, Bode RK, Gershon R, Lai JS. Item banks and their potential applications to health status assessment in diverse populations. *Med Care* 2006;44:S189-S197.
49. Cella D, Lai JS, Nowinski C, et al. Neuro-qol: Brief measures of health-related quality of life for clinical research in neurology. *Neurology* 2012;78:1860-1867.
50. Cella D, Nowinski C, Peterman A, et al. The neurology quality-of-life measurement initiative. *Arch Phys Med Rehabil* 2011;92:S28-S36.
51. Gershon R, Lai J, Bode R, et al. Neuro-QoL: Quality of life item banks for adults with neurological disorders: Item development and calibrations based upon clinical and general population testing. *Qual Life Res* 2012;21:475-486.
52. European Medicines Agency Committee for Medicinal Products for Human Use. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003637.pdf. Accessed 25 July 2017.
53. Perez L, Huang J, Jansky L, et al. Using focus groups to inform the neuro-qol measurement tool: exploring patient-centered, health-related quality of life concepts across neurological conditions. *J Neurosci Nurs* 2007;39:342-353.
54. Miller DM, Bethoux F, Victorson D, et al. Validating Neuro-QoL short forms and targeted scales with people who have multiple sclerosis. *Mult Scler* 2016;22:830-841.
55. Cook KF, Victorson DE, Cella D, Schalet BD, Miller D. Creating meaningful cut-scores for neuro-qol measures of fatigue, physical functioning, and sleep disturbance using standard setting with patients and providers. *Qual Life Res* 2014;24:1-15.
56. Cella D, Yount S, Rothrock N, et al. The patient-reported outcomes measurement information system (PROMIS): progress of an NIH roadmap cooperative group during its first two years. *Med Care* 2007;45:S3-S11.
57. Cella D, Riley W, Stone A, et al. The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol* 2010;63:1179-1194.
58. Cook KF, Jensen SE, Schalet BD, et al. PROMIS® measures of pain, fatigue, negative affect, physical function and social function demonstrate clinical validity across a range of chronic conditions. *J Clin Epidemiol* 2016;73:89-102.
59. Hays R, Bode R, Rothrock N, et al. The impact of next and back buttons on time to complete and measurement reliability in web-based surveys *Qual Life Res*. 2010;19:1181–1184.
60. Northwestern University. HealthMeasures: Data collection tools. Available at: <http://www.healthmeasures.net/resource-center/data-collection-tools>. Accessed 31 July 2017.
61. McHorney CA, Cohen AS. Equating health status measures with item response theory: Illustrations with functional status items. *Med Care* 2000;38:1143-1159.
62. Reeve B, Wyrwich K, Wu A, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013;8:1889-1905.
63. Mokkink LB, Prinsen CAC, Bouter LM, de Vet HCW, Terwee CB. The consensus-based standards for the selection of health measurement instruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther* 2016;20:105-113.
64. Prinsen CAC, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a “core outcome set”—a practical guideline. *Trials* 2016;17:449.
65. Wehrlen L, Krumlauf M, Ness E, Maloof D, Bevans M. Systematic collection of patient reported outcome research data: a checklist for clinical research professionals. *Contemp Clin Trials* 2016;48:21-29.
66. Francis DO, McPheeters ML, Noud M, Penson DF, Feurer ID. Checklist to operationalize measurement characteristics of patient-reported outcome measures. *Syst Rev* 2016;5:129.
67. Multiple Sclerosis Coalition. The use of disease-modifying therapies in multiple sclerosis: Principles and current evidence. National Multiple Sclerosis Society 2017. Available at: http://www.nationalmssociety.org/getmedia/5ca284d3-fc7c-4ba5-b005-ab537d495c3c/DMT_Consensus_MS_Coalition_color. Accessed 4 Sept 2017.
68. Benito-Leon J, Morales JM, Rivera-Navarro J, Mitchell A. A review about the impact of multiple sclerosis on health-related quality of life. *Disabil Rehabil* 2003;25:1291-1303.
69. Rudick RA, Miller D, Clough JD, Gragg LA, Farmer RG. Quality of life in multiple sclerosis. Comparison with inflammatory bowel disease and rheumatoid arthritis. *Arch Neurol* 1992;49:1237-1242.
70. Kurtzke JF. Neurologic impairment in multiple sclerosis and the disability status scale. *Acta Neurol Scand* 1970;46:493-512.
71. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983;33: 1444-1452.
72. Jones CA, Pohar SL, Warren S, Turpin KV, Warren KG. The burden of multiple sclerosis: a community health survey. *Health Qual Life Outcomes* 2008;6:1.
73. McCabe MP, Stokes M, McDonald E. Changes in quality of life and coping among people with multiple sclerosis over a 2 year period. *Psychol Health Med* 2009;14:86-96.
74. Power M, Harper A, Bullinger M. The world health organization WHOQOL-100: tests of the universality of quality of life in 15 different cultural groups worldwide. *Health Psychol* 1999;18:495-505.
75. Flachenecker P, Buckow K, Pugliatti M, et al. Multiple sclerosis registries in europe - results of a systematic survey. *Mult Scler* 2014;20:1523-1532.
76. Marrie RA, Cutter G, Tyry T, Hadjimichael O, Vollmer T. Validation of the narcoms registry: pain assessment. *Mult Scler* 2005;11:338-342.
77. Wu N, Minden SL, Hoaglin DC, Hadden L, Frankel D. Quality of life in people with multiple sclerosis: Data from the sonja slifka longitudinal multiple sclerosis study. *J Health Hum Serv Adm* 2007;30:2332-67.
78. Accelerated Cure Project for MS. Iconquerms. Available at: <https://iconquerms.org/>. Accessed 15 March 2017.

79. Ware J, Jr., Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220-233.
80. Baumstarck K, Pelletier J, Boucekine M, Auquier P. Predictors of quality of life in patients with relapsing-remitting multiple sclerosis: a 2-year longitudinal study. *Rev Neurol (Paris)* 2015;171:173-180.
81. Vermersch P, de Seze J, Delisse B, Lemaire S, Stojkovic T. Quality of life in multiple sclerosis: Influence of interferon-beta 1 a (Avonex) treatment. *Mult Scler* 2002;8:377-381.
82. Lily O, McFadden E, Hensor E, Johnson M, Ford H. Disease-specific quality of life in multiple sclerosis: the effect of disease modifying treatment. *Mult Scler* 2006;12:808-813.
83. Alschuler K, Ehde D, Kratz A, Jensen M. Prevalence of pain and associations with functioning and quality of life in individuals newly-diagnosed with multiple sclerosis. *J Pain* 2016;17:S18-S19.
84. Fernandez-Munoz JJ, Moron-Verdasco A, Cigaran-Mendez M, et al. Disability, quality of life, personality, cognitive and psychological variables associated with fatigue in patients with multiple sclerosis. *Acta Neurol Scand* 2015;132:118-124.
85. Kinkel RP, Laforet G, You X. Disease-related determinants of quality of life 10 years after clinically isolated syndrome. *Int J MS Care* 2015;17:26-34.
86. Baumstarck-Barrau K, Simeoni MC, Reuter F, et al. Cognitive function and quality of life in multiple sclerosis patients: a cross-sectional study. *BMC Neurol* 2011;11:17.
87. Campbell J, Rashid W, Cercignani M, Langdon D. Cognitive impairment among patients with multiple sclerosis: associations with employment and quality of life. *Postgrad Med J* 2017;93:143-147.
88. Samartzis L, Gavala E, Zoukos Y, Aspiotis A, Thomaidis T. Perceived cognitive decline in multiple sclerosis impacts quality of life independently of depression. *Rehabil Res Pract* 2014;2014: Article ID 128751.
89. Marck CH, Jelinek PL, Weiland TJ, et al. Sexual function in multiple sclerosis and associations with demographic, disease and lifestyle characteristics: an international cross-sectional study. *BMC Neurol* 2016;16:210.
90. Schairer LC, Foley FW, Zemon V, et al. The impact of sexual dysfunction on health-related quality of life in people with multiple sclerosis. *Mult Scler* 2014;20:610-616.
91. Vitkova M, Rosenberger J, Krokavcova M, et al. Health-related quality of life in multiple sclerosis patients with bladder, bowel and sexual dysfunction. *Disabil Rehabil* 2014;36:987-992.
92. Nourbakhsh B, Julian L, Waubant E. Fatigue and depression predict quality of life in patients with early multiple sclerosis: a longitudinal study. *Eur J Neurol* 2016;23:1482-1486.
93. Wang JL, Reimer MA, Metz LM, Patten SB. Major depression and quality of life in individuals with multiple sclerosis. *Int J Psychiatry Med*. 2000;30:309-317.
94. Miller DM, Kattan MW, Fu AZ. Health related quality of life assessment in multiple sclerosis. In: Cohen JA, Rudick RA, editors. *Multiple sclerosis therapeutics*. 4th ed. New York: Cambridge University Press; 2011. p. 79-90.
95. Rudick RA, Miller DM. Health-related quality of life in multiple sclerosis: Current evidence, measurement and effects of disease severity and treatment. *CNS Drugs* 2008;22:827-839.
96. Freeman JA, Thompson AJ, Fitzpatrick R, et al. Interferon-beta1b in the treatment of secondary progressive MS: impact on quality of life. *Neurology* 2001;57:1870-1875.
97. Rudick RA, Stuart WH, Calabresi PA, et al. Natalizumab plus interferon beta-1a for relapsing multiple sclerosis. *N Engl J Med* 2006;354:911-923.
98. Kappos L, Polman CH, Freedman MS, et al. Treatment with interferon beta-1b delays conversion to clinically definite and McDonald MS in patients with clinically isolated syndromes. *Neurology* 2006;67:1242-1249.
99. Panitch H, Miller A, Paty D, Weinschenker B, North American Study Group on Interferon beta-1b in Secondary Progressive MS. Interferon beta-1b in secondary progressive MS: results from a 3-year controlled study. *Neurology* 2004;63:1788-1795.
100. Cohen JA, Cutter GR, Fischer JS, et al. Benefit of interferon beta-1a on MSFC progression in secondary progressive MS. *Neurology* 2002;59:679-687.
101. Jongen PJ. Health-related quality of life in patients with multiple sclerosis: Impact of disease-modifying drugs. *CNS Drugs* 2017;31:585-602.
102. Calabresi PA, Kieseier BC, Arnold DL, et al. Pegylated interferon beta-1a for relapsing-remitting multiple sclerosis (advance): A randomised, phase 3, double-blind study. *Lancet Neurol* 2014;13:657-665.
103. Kieseier BC, Arnold DL, Balcer LJ, et al. Peginterferon beta-1a in multiple sclerosis: 2-year results from advance. *Mult Scler* 2015;21:1025-1035.
104. Gold R, Giovannoni G, Selmaj K, et al. Daclizumab high-yield process in relapsing-remitting multiple sclerosis (select): a randomised, double-blind, placebo-controlled trial. *Lancet* 2013;381:2167-2175.
105. Gold R, Kappos L, Arnold DL, et al. Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. *N Engl J Med* 2012;367:1098-1107.
106. Newsome SD, Guo S, Altincatal A, et al. Impact of peginterferon beta-1a and disease factors on quality of life in multiple sclerosis. *Mult Scler Relat Disord* 2015;4:350-357.
107. Phillips G, Guo S, Bender R, et al. Assessing the impact of multiple sclerosis disease activity and daclizumab hyp treatment on patient-reported outcomes: results from the select trial. *Mult Scler Relat Disord* 2016;6:66-72.
108. Kappos L, Wiendl H, Selmaj K, et al. Daclizumab hyp versus interferon beta-1a in relapsing multiple sclerosis. *N Engl J Med* 2015;373:1418-1428.
109. Confavreux C, O'Connor P, Comi G, et al. Oral teriflunomide for patients with relapsing multiple sclerosis (tower): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Neurol* 2014;13:247-256.
110. O'Connor P, Wolinsky JS, Confavreux C, et al. Randomized trial of oral teriflunomide for relapsing multiple sclerosis. *N Engl J Med* 2011;365:1293-1303.
111. Calabresi PA, Radue EW, Goodin D, et al. Safety and efficacy of fingolimod in patients with relapsing-remitting multiple sclerosis (FREEDOMS II): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Neurol* 2014;13:545-556.
112. Kita M, Fox RJ, Phillips JT, et al. Effects of BG-12 (dimethyl fumarate) on health-related quality of life in patients with relapsing-remitting multiple sclerosis: findings from the confirm study. *Mult Scler* 2014;20:253-257.
113. Fisk JD, Ritvo PG, Ross L, et al. Measuring the functional impact of fatigue: Initial validation of the fatigue impact scale. *Clin Infect Dis* 1994;18(Suppl. 1):S79-S83.
114. Cohen JA, Barkhof F, Comi G, et al. Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. *N Engl J Med* 2010;362:402-415.
115. Arroyo González R, Kita M, Crayton H, et al. Alemtuzumab improves quality-of-life outcomes compared with subcutaneous interferon beta-1a in patients with active relapsing-remitting multiple sclerosis. *Mult Scler* 2017;23:1367-1376.
116. Banta D, Kristensen FB, Jonsson E. A history of health technology assessment at the European level. *Int J Technol Assess Health Care* 2009;25:68-73.
117. Hailey D. Development of the international network of agencies for health technology assessment. *Int J Technol Assess Health Care* 2009;25:24-27.

118. Nielsen CP, Funch TM, Kristensen FB. Health technology assessment: research trends and future priorities in europe. *J Health Serv Res Policy* 2011;16:6-15.
119. Patient Centered Outcomes Research Institute. Research we support: National priorities and research agenda. Available at: <http://www.pcori.org/research-results/research-we-support>. Accessed 25 March 2017.
120. Patient Centered Outcomes Research Institute. Treatment of multiple sclerosis: Meeting summary. PCORI : Prioritizing Comparative Effectiveness Research Questions for the Treatment of Multiple Sclerosis: A Stakeholder Workshop [serial on the Internet]. Available at: <http://www.pcori.org/sites/default/files/PCORI-Prioritizing-CER-Questions-Treatment-of-MS-Workshop-Meeting-Summary-040215.pdf>. Accessed 4 Sept 2017.
121. Cohen JA, Cutter GR, Fischer JS, et al. Use of the multiple sclerosis functional composite as an outcome measure in a phase 3 clinical trial. *Arch Neurol* 2001;58:961-967.
122. Cutter GR, Baier ML, Rudick RA, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain* 1999;122:871-882.
123. Fischer JS, Rudick RA, Cutter GR, Reingold SC. The multiple sclerosis functional composite measure (MSFC): an integrated approach to ms clinical outcomes assessment. *Mult Scler* 1999;5:244-250.
124. Rudick RA, Cutter G, Baier M, et al. Use of the multiple sclerosis functional composite in predicting disability in relapsing MS. *Neurology* 2001;56:1324-1330.
125. Dadalti FY. Why some of us do not like the expression "no evidence of disease activity" (NEDA) in multiple sclerosis. *Mult Scler Relat Disord* 2015;4:383-384.
126. Giovannoni G, Tumer B, Gnanapavan S, et al. Is it time to target no evident disease activity (NEDA) in multiple sclerosis? *Mult Scler Relat Disord* 2015;4:329-333.
127. Ziemssen T, Derfuss T, de Stefano N, et al. Optimizing treatment success in multiple sclerosis. *J Neurol* 2016;263:1053-1065.
128. Institute of Medicine. *Best care at lower prices: The path to continuously learning health care in america*. Washington, DC: The National Academies Press; 2013.
129. Bermel RA, Mowry EM, Krupp L, et al. Multiple sclerosis partners advancing technology and health solutions (MS paths): initial launch experience. American Academy of Neurology; Boston, MA, 2017.
130. NIH Collaboratory. Rethinking clinical trials; report on working groups. Available at: <https://www.nihcollaboratory.org/cores/Pages/PRO.aspx>. Accessed 4 Sept 2017.
131. Centers for Medicare & Medicaid Services. CMS quality measure development plan: Supporting the transition to the merit-based incentive payment system (MIPS) and alternative payment models (APMS). Baltimore, MD, 2016.
132. Schlesinger M, Grob R, Shaller D. Using patient-reported information to improve clinical practice. *Health Serv Res* 2015;50(Suppl. 2):2116-2154.