Outcome Measures in Neurological Physical Therapy Practice: Part I. Making Sound Decisions

Kirsten Potter, PT, DPT, MS, NCS, George D. Fulk, PT, PhD, Yasser Salem, PT, PhD, NCS, PCS, and Jane Sullivan, PT, DHS

Standardized outcome measures (OMs) are a vital part of evidence-based practice. Despite the recognition of the importance of OMs, recent evidence suggests that the use of OMs in clinical practice is limited. Selecting the most appropriate OM enhances clinical practice by (1) identifying and quantifying body function and structure limitations; (2) formulating the evaluation, diagnosis, and prognosis; (3) informing the plan of care; and (4) helping to evaluate the success of physical therapy interventions. This article (Part I) is the first of a 2-part series on the process of selecting OMs in neurological clinical practice. We introduce a decision-making framework to guide the selection of OMs and discuss 6 main factors—what to measure, the purpose of the measure, the type of measure, patient and clinic factors, psychometric factors, and feasibility—that should be considered when selecting OMs for clinical use. The framework will then be applied to a patient case in Part II of the series (see the article "Outcome Measures in Neurological Physical Therapy Practice: Part II. A Patient-Centered Process" in this issue).

Key words: *examination, measurement, neurological physical therapist practice, outcome measures*

(*JNPT* 2011;35: 57–64)

INTRODUCTION

Standardized outcome measures (OMs) are an essential tool for evidence-based practice. Outcome measures generate scores that are intended to quantify a patient's performance or health status based on standardized evaluation protocols or close-ended questions.[1] Outcome measures differ from certain measures used in physical therapist (PT) practice (eg, observation of posture and movement, manual muscle test, and goniometry) in that they are intended to track change in performance or health status. The psychometric properties of many OMs have been evaluated in a given target population, providing PTs with the current best evidence for use when making clinical decisions. Outcome measures are often used to evaluate change in a patient's status following therapeutic intervention. When used at the initial examination or early in the episode of care, OMs enhance and support clinical decision making by helping therapists develop the plan of care and set goals.[1,2] Subsequently, periodic use of OMs may lead to modification of the patient's plan of care and, when used at the end of an episode of care, allow determination of change over time and overall effectiveness of the intervention. Importantly, they also assist when communicating with patients,[3,4] other health care professionals, and third-party payers about change in performance or health status.[1,2] In addition to evaluating change, some OMs can be used to diagnose a patient's condition and/or predict future status; however, these uses of OMs are not the focus of this article.

Despite recent evidence-based practice initiatives and the need for accountability that have highlighted the need to use OMs,[5,6] recent studies show that their use in clinical practice remains limited.[1,7] Various challenges are associated with the selection and use of OMs. Surveyed PTs indicate that barriers include time constraints, difficulty for patients to complete, lack of equipment, and lack of knowledge regarding OMs.[1,7,8] A limited understanding of how to select and apply the best OM has been reported to be a barrier to their use in clinical practice.[9]

Decision-making frameworks help to guide PTs through the patient management process. There are a variety of frameworks that can be used to assist in clinical decision making in neurological PT practice.[10–13] Although these are useful for a number of purposes including diagnosis, evaluation, task analysis, progression, and comprehensive patient management, they do not provide specific guidance when selecting OMs. This article (Part I) is the first of a 2-part series on the processes of selecting OMs in neurological clinical practice. The purposes of this article are to introduce a clinical decision-making framework to guide the examination and OM selection process, and to discuss factors that are important to consider when selecting OMs. The framework will then be applied to a patient case in Part II of the series.

Department of Physical Therapy and Human Movement Sciences (K.P., J.S.), Northwestern University, Feinberg School of Medicine, Chicago, Illinois; Department of Physical Therapy (G.D.F.), Clarkson University, Potsdam, New York; and Division of Physical Therapy (Y.S.), Long Island University, Brooklyn, New York.

No funding received for this work was received from any of the following organizations: National Institutes of Health; Wellcome Trust; or the Howard Hughes Medical Institute.

This article derives from the work developed for the Neurology Section Regional Continuing Education Course: Neurologic Practice Essentials: A Measurement Toolbox.

Correspondence: Kirsten Potter, PT, DPT, MS, NCS, Department of Physical Therapy and Human Movement Sciences, Northwestern University, Feinberg School of Medicine, 645 N Michigan Ave, Ste 1100, Chicago, IL 60611 (k-potter@northwestern.edu).

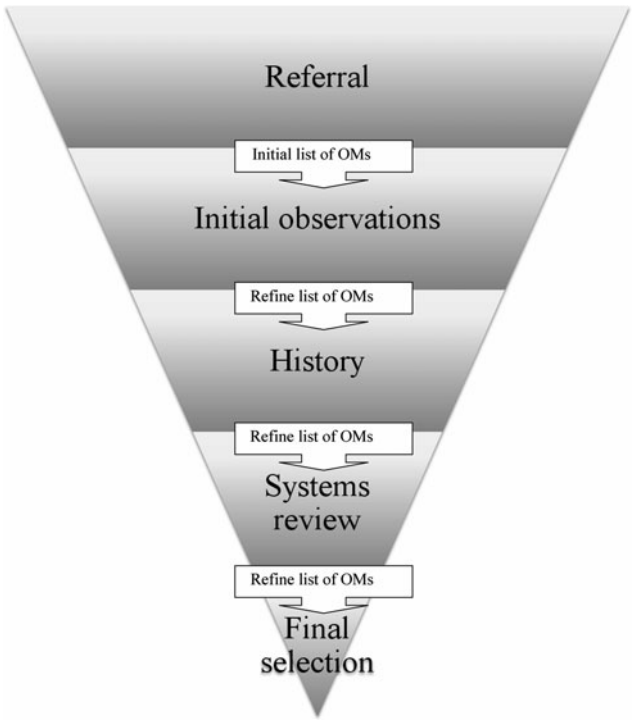
ISSN: 1557-0576/11/3502-0057
DOI: 10.1097/NPT.0b013e318219a51a

OVERVIEW OF THE FRAMEWORK

The framework (Figure 1) we propose uses a deductive reasoning process to guide the process of examination and OM selection. The framework begins with a patient's referral to physical therapy, when the PT receives the patient's medical diagnosis and reason for referral. This information is used to generate an initial list (potentially quite broad and imprecise) of OMs, which may be appropriate for use, based on the PT's knowledge of the patient's health condition, experience working with patients with similar conditions, and previous experience with OMs. Gathering the patient's medical history, through a review of the patient's chart or during the interview or both, further narrows the list of possible OMs. Upon meeting the patient, the PT observes the patient's movement, formulating initial hypotheses about impairments underlying the movement problem. This information is used to refine the list of possible OMs, resulting in the elimination of certain OMs or identifying new OMs or both for further consideration.

The next step in the process, the systems review/screening, identifies those areas needing further examination. The PT further refines the list of OMs, selecting those most useful to understanding and quantifying the patient's limitations and abilities. Throughout this process, as the PT refines the list of OMs and progresses toward the final selection, as illustrated in Figure 1, 6 main factors need to be considered at each step. These factors are as follows: what to measure, purpose of the measure, type of measure, patient and clinic factors, psychometric factors, and feasibility (Table 1). This framework can be used by clinicians working in any setting, with patients with any type of health condition. In addition, educators can use the framework to teach students how to select appropriate OMs.

Figure 1. Clinical decision-making framework: the steps in selecting outcome measures. As the physical therapist proceeds through the examination process, she or he refines and narrows the list of possible OMs, leading to the final selection of OM(s).



WHAT TO MEASURE?

Two conceptual frameworks are used to assist PTs when deciding what to measure: the World Health Organization's International Classification of Functioning, Disability, and Health (ICF)[14,15] and the *Guide to Physical Therapist Practice* (*Guide*).[16] The *Guide* identifies 24 categories of tests and measures (eg, Arousal, Attention, and Cognition; Motor Function; Sensory Integrity; and Self-Care and Home Management) that broadly outline what constructs to measure. Within each category, specific aspects of a construct that could be measured are outlined. For example, under Motor Function, the *Guide* suggests assessing dexterity, coordination, and agility using coordination screens, motor impairment tests, and motor proficiency tests. However, specific OMs are not mentioned. The *Interactive Catalog of Tests and Measures*[17] names specific tests/measures (including OMs with psychometric data) relevant to each category of test and measure. However, the *Guide* makes no recommendations for the use of 1 OM over another. Thus, although the *Guide* and the *Interactive Catalog of Tests and Measures* assist with determining what to measure, the selection of OM(s) is left to the judgment of the PT.

The ICF consists of 2 interactive parts (Table 2).[14,15] Part 1, Functioning and Disability, addresses factors intrinsic to the individual and has 2 components: Body Function and Structure, and Activity and Participation. Part 2, Contextual Factors, includes matters extrinsic to the individual: Environmental and Personal Factors. The ICF model offers several benefits to PTs and other health care providers.[18–20] The ICF emphasizes the importance of viewing the patient holistically by considering the interactive effects of the patient's health condition. Thus, PTs should consider using OMs across *all* levels of the ICF, including participation, to gain a broad understanding of the impact of a patient's health condition.

Table 1. Factors to Consider When Selecting Outcome Measures for Clinical Practice

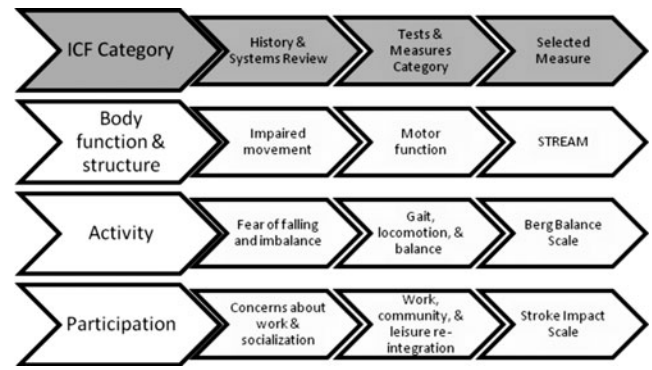
What to measure	Body function/structure, activity, participation
Purpose of measure	Discriminative
	Predictive
	Evaluative
Type of measure	Disease-specific or generic
	Self-report or performance-based
Patient and clinic factors	Patient recovery and ability
	Patient goals
	Clinic requirements
Psychometric factors	Reliability
	Validity
	Diagnostic accuracy
	Responsiveness
Feasibility	Time, space, and equipment
	Training required
	Cost
	Respondent burden
	Culture and language
	Proprietary issues

Table 2. Overview of the International Classification of Functioning, Disability, and Health[14,15]

Components	Definitions (Related Terms)	Examples
Part I: Functioning and disability		
Body function	Anatomical status (impairment of body function)	Joint contracture Muscle atrophy
Body structure	Physiological status (impairment of body structure)	Altered consciousness Impaired muscle timing and sequencing
Activity	Execution of task (activity limitation)	Inability to walk Difficulty with bathing and dressing
Participation	Involvement in life situation (participation restriction)	Inability to work Altered sexual functioning
Part II: Contextual factors		
Environmental factors	Physical, social, and attitudinal environment	Access to health care services Stairs in home
Personal factors	Personal factors that play a role in health	Age Race Education

The *Guide* and the ICF are used collectively to enhance patient management in the selection of OMs, as shown in Figure 2. Consideration of the 3 levels of the ICF enhances comprehensive patient management. During the history and systems review (components of the *Guide*'*s* patient management process), the PT identifies potential constructs that may be important to measure. This leads to the identification of relevant categories of tests and measures, which help to identify more specifically what to measure, informing the selection of appropriate OMs.

Figure 2. Integration of the ICF and the *Guide to Physical Therapist Practice.* The ICF levels (body function/structure, activity, participation) are integrated with the components of the physical therapy examination (history, systems review, and tests and measures), leading to the identification of outcome measures. Abbreviations: ICF, International Classification of Functioning, Disability, and Health; STREAM, Stroke Rehabilitation Assessment of Movement.



PURPOSE OF THE MEASUREMENT

Outcome measures can be classified in various ways, one of which pertains to the purpose of the OM. Outcome measures are generally designed to fulfill 1 of 3 specific purposes: to discriminate, predict, and/or evaluate.[21–23] In many cases, a single OM can serve all 3 purposes, but in other cases an OM was developed to serve only 1 purpose. An OM developed for 1 purpose may not adequately serve others. Hence, it is important to use an OM for the purpose(s) for which it was created and validated.

Discriminative OMs distinguish between individuals or groups on the basis of a particular characteristic. They also separate patients into groups on the basis of predefined categories. Discriminative OMs generally have fewer response choices, limiting their ability to detect small changes after intervention.[24] For example, the Performance-Oriented Mobility Assessment,[25] for measuring balance and gait, uses a 2- to 3-point ordinal scale (eg, Item 7: "Stand with Eyes Closed" is scored as follows: 0 = unsteady or 1 = steady). While the Performance-Oriented Mobility Assessment has excellent discriminative ability (eg, it can distinguish between elders with vs without a history of falls), which is useful in identifying patients who might benefit from PT intervention, this type of scale has limited ability to detect change.[26] Predictive measures are intended to forecast future status and can be used to identify risk, determine a prognosis, and plan discharge. For example, the Berg Balance Scale (BBS) is often used to predict risk of falling.[27–29] Evaluative measures help to determine the effectiveness of an intervention by evaluating change over time. The BBS also serves this purpose, as it is able to detect change in patients with neurological conditions.[30,31] Gait speed is another example of an OM that serves multiple purposes. Gait speed measured over a short distance can be used to categorize patients as home or community ambulators,[32] predict discharge destination from the hospital,[33] and evaluate change in walking ability.[34] Timed tests (eg, the 10-Meter Walk Test) are particularly useful as evaluative measures because they can reveal small changes in ability.

TYPE OF MEASURE

Outcome measures can also be classified according to focus (generic or disease-specific) and mode of administration

(performance-based or self-report). Generic OMs are intended for use across all patient populations, regardless of specific health condition, while disease-specific OMs are designed for use only with a specific patient population. For example, the 36-Item Short Form Health Survey[35] is a generic measure of health-related quality of life that has been validated for use with the general population, while the Parkinson's Disease Questionnaire–39 (PDQ-39)[36] is a health-related quality-of-life measure designed and validated specifically for people with Parkinson's disease. Generic and disease-specific OMs that have published normative values are useful for comparison and goal-setting. In general, generic OMs are more likely to have normative values and can be useful for programmatic evaluation. For example, the Functional Independence Measure has been used to evaluate outcomes obtained across inpatient rehabilitation and skilled nursing facilities.[37] However, a growing number of disease-specific OMs also have normative values, allowing the comparison of an individual patient with the patient population. Disease-specific OMs are likely to measure constructs that are more meaningful to the patient and may be more beneficial when setting goals and determining a plan of care because the items contained therein relate to the disease and its impact on the person. For example, the PDQ-39 asks specifically whether the patient is embarrassed in public due to having Parkinson's disease. Other factors to consider when choosing between generic or disease-specific OMs are outlined in Table 3.[38–40]

Performance-based OMs are used to assess a patient's performance on a set of predetermined activities in the specific environment in which the test is conducted. A performance-based OM provides insight into a patient's actual capability at the point in time the test was administered. A limitation of performance-based OMs is that a patient's capability measured in the clinic may not reflect actual performance in his or her home and community.[41,42] Self-report OMs provide information on the patient's opinions/perceptions of the impact of the health condition. They may also provide information on the patient's perception of activities performed in his or her home and community, which cannot be directly observed in the clinic.[43] A limitation of self-report measures is that patients may over- or underestimate their abilities and/or report what they believe the rater wants to hear.

Available evidence indicates that there is a moderate association between performance-based and self-report measures.[44–47] However, each offers distinct and complementary information,[44,48] suggesting the need for inclusion of both to obtain comprehensive information about function and disability. For example, the BBS is a performance-based measure of functional balance that assesses a person's ability to maintain their balance while they perform 14 different tasks, while the Activities-specific Balance Confidence Scale[49] is a self-report OM that asks patients to rate their confidence in their ability to maintain their balance when performing each of 16 activities in their home and community. By administering both OMs, it is possible for a PT to ascertain whether there is a discrepancy between the patient's actual versus perceived capability related to balance. Important factors when considering performance-based versus self-report OMs are outlined in Table 4.[50–53]

PATIENT AND CLINIC FACTORS

Several factors related to the patient and clinical setting must be considered when selecting OMs, as discussed briefly in the following sections.

Patient Factors

It is of primary importance that the selected OM closely match the patient's goals and the planned intervention.[54,55] Outcome measures can serve to motivate patients by providing feedback related to goal attainment.[56] Selecting OMs across ICF levels provides a means to include measures that are related to the intervention and are clinically important and relevant to the patient.

The patient's stage of recovery impacts OM selection. Early in the recovery phase, measures of body structure/function are likely to provide the best markers of prognosis.[57] Outcome measures that assess body structure/function and activity are more likely to be meaningful to the patient and

Table 3. Comparison of Generic Versus Disease-Specific Measures

	Generic	Disease Specific
Purpose/information gained	Intended for use with all individuals Overall impact of health condition Global and long-term health outcomes	Gather information about the impact of a specific health condition
Population	General population (healthy and nonhealthy) Applicable across individuals with differing health conditions	Used for patients with a specific health condition
Benefits	Often have normative values for groups of individuals (eg, those with a specific health condition or characteristic) Can compare across populations or patient groups	Provides insight into the relationships among body function/structure impairments, activity limitations, and participation restrictions
Limitations	Ceiling and floor effects more likely	Does not allow comparisons across different groups of patients
Examples	Functional Independence Measure Short form-36	Stroke Impact Scale Parkinson's Disease Questionnaire–39

Table 4. Comparison of Performance-Based Versus Self-Report Measures

	Performance Based	Self-report
Purpose and information gained	Assess actual performance in a particular environment at a specific point in time Determine current level of function	Gather information on patient's opinions/perceptions of the impact of the health condition
Data collection methods used	Timing Measuring distance Observation of performance Rating level of independence or difficulty	Interview Questionnaire
Benefits	Allows observation of task performance Less likely to have missing data	Can assess constructs not observed during performance (eg, fear of falling) Can assess activities not observed clinically
Limitations	Equipment required Performance may not generalize to other environments Subject to fatigue, motivation, and learning effects Sources of inaccuracy (eg, improper use of equipment, intrasubject performance variability)	Cognitive and communication deficits may interfere Language and educational level may interfere Patient may miss or misinterpret items or over-/underestimate ability
Examples	Functional Independence Measure 10-Meter Walk Test Timed Up and Go Test	Activities-specific Balance Confidence Scale Modified Fatigue Impact Scale Motor Activity Log

PT and are often more directly linked to the intervention being provided during the early stage of recovery. However, later in the recovery process, measures of participation are likely to be more meaningful to the patient and PT. For example, 6 months after a spinal cord injury in an outpatient setting, the Craig Handicap Assessment and Reporting Technique[58] (which measures disability after spinal cord injury) provides information about the impact of the spinal cord injury on mobility, physical independence, occupation, social integration, and economic self-sufficiency. This information is likely to be more useful to the clinician than the information about muscle strength. However, early after spinal cord injury, knowledge of upper extremity strength is likely to be more useful than social integration for treatment planning and goal-setting.

The patient's capability should also be considered when selecting an OM. The items on the OM should not be too difficult or too easy for the patient to perform. If test items are at either of these extremes, then the OM is likely to exhibit a floor effect or ceiling effect. Floor effects limit the ability of the OM to detect decreased performance beyond a certain level, while ceiling effects limit the ability of the OM to detect increased performance beyond a certain level (assuming that a higher score indicates better performance on the given measure).[53] For example, if a patient has difficulty with sitting balance and requires assistance to sit on the edge of the bed, a measure of balance such as the Functional Gait Assessment[59] is not appropriate; however, the BBS or the Postural Assessment Scale for Stroke Patients[23] would be appropriate.

Clinic Requirements

The requirements of the clinical setting are important to consider when selecting OMs. Some settings require that certain OMs be administered with all patients (eg, the Functional Independence Measure in acute inpatient rehabilitation units). Although this requirement for the use of specific measures can be useful for programmatic evaluation, it can have implications for the clinician (eg, feasibility in terms of time available to administer other OMs).

PSYCHOMETRIC FACTORS

In order for an OM to provide information that can be used to guide clinical decision making, it must exhibit sound psychometric properties[53,60] including reliability and validity, and the related properties of diagnostic accuracy and responsiveness (ie, the ability of an OM to evaluate change).[22,53,60–62] It is beyond the scope of this article to provide an in-depth review of psychometric properties. There are many excellent resources available for readers who desire more information on this topic.[22,53,60,61] Here we focus on 2 important psychometric properties to consider when selecting OMs to evaluate change, the minimal detectable change (MDC) and minimal clinically important difference (MCID).[63]

Measurement error is particularly important to consider when selecting an OM to evaluate change over time. When taking a measurement at 2 different times, one wants to be confident that the change in the OM scores resulted from real change, not from measurement error. The MDC, also referred to as the minimal detectable difference, is the minimal amount of change on the OM that is necessary to exceed error and variability in that measure.[53,63] The MDC is considered to be the smallest amount of change in an OM that is necessary for the change to be considered a true change.

The MDC is useful clinically for evaluating change because it is presented in the units of the OM. For example, Stevenson[64] determined an MDC of 7 points for the BBS in people with stroke undergoing inpatient rehabilitation. A clinician can use this MDC value to interpret change scores. For example, consider a change from an initial BBS score of 36 to a score of 40 three weeks later. This 4-point change in the BBS did not exceed measurement error (MDC of 7 points). The clinician would conclude that the patient's functional balance, as measured by the BBS, had not changed. This may lead

the clinician to reevaluate the plan of care as the interventions selected may not be effective for improving functional balance in this patient.

The interpretation of MDC values must be done within the context of the studies from which they were developed. Minimal detectable change values will vary depending on the characteristics of the patients and the raters. For example, various authors have reported the MDC of gait speed in people with stroke. These values range from 0.07 to 0.36 m/s.[34,65–67] The disparity in these values is likely due to differences between the acuity of the subjects, need for physical assistance or an assistive device when walking, the walking surface, method of measuring gait speed, the rater's experience in the studies, and possible changes in the subject's performance in between measurements. When selecting an OM with a reported MDC, clinicians should strive to select one such that the characteristics of their patient are similar to those of the subjects in the study from which the MDC was determined.

An important limitation of the MDC is that it only estimates the threshold for true change; it does not provide information about whether the change was clinically important. The MCID value of an OM can be used to assess whether or not a change in an OM was meaningful.[53,63] The MCID is defined as the smallest difference in a measure that is considered worthwhile or important. The MCID is often based on the relationship between the OM of interest and some independent, external anchor or criterion to determine clinically important change.[68,69] Change scores in an OM that exceed the MCID value indicate that the patient's improvement was clinically important on the basis of some external criterion. For example, Lang et al[70] determined the MCID of various upper extremity measures, including the Motor Activity Log (MAL), in people in the acute stage of recovery poststroke using patients' global perception of change in upper extremity function as the external criterion. They found an MCID of 1.0 on the MAL. If a patient's score on the MAL changed from 1.5 to 3.0, this 1.5-point improvement exceeds the MCID, indicating that the amount of change was likely clinically important.

The limitations of the MCID are similar to those of the MDC. MCID scores have limited generalizability,[63] so they should be used only with patients having characteristics similar to those of the subjects for which MCID were reported. A patient's initial OM score may impact the chance of exceeding the MCID. Patients with lower scores or more severe disability may have an increased chance of exceeding the MCID than patients who initially score high or have less disability (note that this pertains to OMs where a higher score indicates better performance).[63] Most studies that estimate MCIDs report values pertaining to improvement (ie, the MCID value applies to a positive change in the patient's status). It should not be assumed that the same values apply when there is deterioration in the patient's condition (ie, a worse score on reassessment), as may occur in patients with progressive diseases. This is illustrated by Nilsagard et al,[71] who found 2 different change scores representing clinically meaningful improvement (+ 31%) and deterioration (− 24%) on the Timed Up and Go Test in patients with multiple sclerosis. The reader is referred to the article by Beninato and Portney[72] in this issue for a more in-depth examination of issues related to responsiveness.

FEASIBILITY

When determining whether an OM is feasible to use, various factors, including time, space and equipment requirements, training required to administer and interpret the OM, cost to obtain and use the OM, proprietary restrictions, respondent burden, and the patient's language and culture, should be considered. Some OMs may provide valuable information, but because of the time required to complete, the feasibility for use may be limited. For example, the Balance Evaluation Systems Test (BESTest),[73] which was developed to assist in differentiating between possible causes of impaired postural control, takes approximately 35 minutes to administer. A busy clinician may not have the time to include this OM when first examining a patient. However, the mini-BESTest[74] was developed from a subset of items from the BESTest; this test can be administered in 10 to 15 minutes, which may be more feasible for many clinicians.

In some settings, space may be an important limitation to the use of particular OMs. A 6-minute walk test provides useful information regarding a patient's walking endurance, but in a home care setting there may not be ample space to perform this test. Measuring gait speed over a 5-m distance may be more feasible; however, if the PT is interested in capturing walking endurance, then it will be necessary to consider whether capturing walking speed represents a viable alternative. Kinematic and kinetic evaluation of gait using motion analysis systems and force plates can be useful in identifying the potential causes of gait abnormalities and measuring change over time. However, a great deal of training and expertise is required to administer and interpret these tests. Some OMs may place a physical or emotional burden on the patient. For example, quality-of-life measures may include items pertaining to sexuality and sexual functioning that some patients may find embarrassing or emotionally difficult to answer.

The clinician must be sensitive to the patient's culture and language, particularly when selecting self-report measures. It is important that an OM be culturally and linguistically appropriate. The PDQ-39[36] has been translated and tested in more than 40 different languages; hence, it may have benefit to PTs working with patients with broad cultural diversity. Clinicians may also need to consider proprietary issues when selecting OMs. Some OMs are protected by copyright and must be purchased.

SUMMARY

The use of standardized OMs is essential for physical therapy practice. They support and enhance practice in a variety of ways. Outcome measures inform patient care by identifying and quantifying body function and structure impairments, and activity and participation limitations; they assist with formulating the evaluation, diagnosis, and prognosis, and help to evaluate the success of physical therapy intervention. Outcome measures can also be used to provide motivation, inform patients and caregivers about the outcomes of PT, and justify PT services. When OMs are administered across groups of patients, they can be used to evaluate clinical programs and for quality assurance.

Selecting the appropriate OMs to use with a patient is essential. If the incorrect OMs are used, it may be difficult to

determine the most appropriate plan of care, evaluate the effectiveness of the intervention, document change, and receive reimbursement. On the contrary, selecting appropriate OMs can enhance practice by helping to set measurable and achievable goals, motivate the patient, determine the effectiveness of the intervention, and ascertain whether the plan of care needs to be modified.

In this article, we presented a clinical decision-making framework to guide the PT examination process, including the selection of OMs. By evaluating the 6 factors (what to measure, the purpose of the measurement, type of measurement, patient/clinic factors, psychometric factors, and feasibility) at each step of the framework, PTs can select OMs that will enhance and guide their practice. Undoubtedly, the process is much more complex than illustrated in Figures 1 and 2, both of which imply a clear-cut, linear process. In reality, the PT must consider and integrate multiple pieces of information in an efficient and effective manner to narrow the selection of tests and measures, an undertaking that is particularly challenging for the novice clinician. A common dilemma is the desire to comprehensively evaluate a patient, a goal that is prevented by limited time. A deductive reasoning process (as outlined in Figure 1) can assist with narrowing the list of possible OMs to help meet the time demands commonly faced by practicing PTs. In addition, facility requirements necessitating the administration of certain OMs may limit a PT's ability to use other OMs that provide additionally useful information. In Part II of this series,[75] the framework described in this article is used for a patient poststroke, wherein various OMs measuring the same constructs are compared and contrasted resulting in the selection of one of the OMs.

REFERENCES

1. Jette DU, Halbert J, Iverson C, Miceli E, Shah P. Use of standardized outcome measures in physical therapist practice: perceptions and applications. *Phys Ther.* 2009;89:125-135.
2. Haigh R, Tennant A, Biering-Sorensen F, et al. The use of outcome measures in physical medicine and rehabilitation within Europe. *J Rehabil Med.* 2001;33:273-278.
3. Patrick DL, Chiang YP. Measurement of health outcomes in treatment effectiveness evaluations: conceptual and methodological challenges. *Med Care.* 2000;38(9 suppl):II14- II25.
4. Wade DT. Assessment, measurement and data collection tools. *Clin Rehabil.* 2004;18:233-237.
5. Jette AM, Haley SM. Contemporary measurement techniques for rehabilitation outcomes assessment. *J Rehabil Med.* 2005;37:339-345.
6. Herbert R, Jamtvedt G, Mead J, et al. Outcome measures measure outcomes, not effects of intervention [see comment]. *Aust J Physiother.* 2005;51:3-4.
7. Van Peppen RP, Maissan FJF, Van Genderen FR, Van Dolder R, Van Meeteren NLU. Outcome measures in physiotherapy management of patients with stroke: a survey into self-reported use, and barriers to and facilitators for use. *Physiother Res Int.* 2008;13:255-270.
8. Kay TM, Myers AM, Huijbregts MP. How far have we come since 1992? A comparative survey of physiotherapists' use of outcome measures. *Physiother Can.* 2001;53:268-281.
9. Huijbregts MP, Myers AM, Kay TM, Gavin TS. Systematic outcome measurement in clinical practice: challenges experienced by physiotherapists. *Physiother Can.* 2002;54:25-36.
10. Schenkman M, Bliss ST, Day L, Kemppainen S, Morse J, Pratt J. Multisystem model for management of neurologically impaired adults—an update and illustrative case. *Neurol Rep.* 1999;23:145-157.
11. Schenkman M, Deutsch JE, Gill-Body KM. An integrated framework for decision making in neurologic physical therapist practice. *Phys Ther.* 2006;86:1681-1702.
12. Scheets PL, Sahrmann SA, Norton BJ. Use of movement system diagnoses in the management of patients with neuromuscular conditions: a multiple-patient case report. *Phys Ther.* 2007;87:654-669.
13. Hedman LD, Rogers MW, Hanke TA. Neurologic professional education: linking the foundation science of motor control with physical therapy interventions for movement dysfunction. *Neurol Rep.* 1996;20:9-13.
14. Towards a common language for functioning, disability and health: ICF. 2002 http://www.who.int/classifications/icf/training/icfbeginnersguide.pdf. Accessed October 12, 2009.
15. International Classification of Functioning, Disability and Health: Introduction. 2001. http://whqlibdoc.who.int/publications/2001/9241545429.pdf. Accessed October 12, 2009.
16. Guide to physical therapist practice. *Phys Ther.* 2001;81:12-746.
17. *Interactive Guide to Physical Therapist Practice, with Catalog of Tests and Measures* [computer program]. Version 1.1. Alexandria, VA: American Physical Therapy Association; 2003.
18. Schepers VP, Ketelaar M, van de Port IG, Visser-Meily JM, Lindeman E. Comparing contents of functional outcome measures in stroke rehabilitation using the International Classification of Functioning, Disability and Health. *Disabil Rehabil.* 2007;29:221-230.
19. Stucki G, Ewert T, Cieza A. Value and application of the ICF in rehabilitation medicine. *Disabil Rehabil.* 2003;25:628-634.
20. Bilbao A, Kennedy C, Chatterji S, Ustün B, Barquero JL, Barth JT. The ICF: applications of the WHO model of functioning, disability and health to brain injury rehabilitation. *Neurorehabilitation.* 2003;18:239-250.
21. Brock KA, Goldie PA, Greenwood KM. Evaluating the effectiveness of stroke rehabilitation: choosing a discriminative measure. *Arch Phys Med Rehabil.* 2002;83:92-99.
22. Barak S, Duncan PW. Issues in selecting outcome measures to assess functional recovery after stroke. *NeuroRx.* 2006;3:505-524.
23. Hsieh CL, Sheu CF, Hsueh IP, Wang CH. Trunk control as an early predictor of comprehensive activities of daily living function in stroke patients. *Stroke.* 2002;33:2626-2630.
24. Roach KE. Measurement of health outcomes: reliability, validity, and responsiveness. *J Pros Orthot.* 2006;18:8-12.
25. Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. *J Am Geriatr Soc.* 1986;34:119-126.
26. Lin MR, Hwang HF, Hu MH, Wu HD, Wang YW, Huang FC. Psychometric comparisons of the Timed Up and Go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. *J Am Geriatr Soc.* 2004;52:1343-1348.
27. Whitney S, Wrisley D, Furman J. Concurrent validity of the Berg Balance Scale and the Dynamic Gait Index in people with vestibular dysfunction. *Physiother Res Int.* 2003;8:178-186.
28. Wrisley DM, Walker ML, Echternach JL, Strasnick B. Reliability of the dynamic gait index in people with vestibular disorders. *Arch Phys Med Rehabil.* 2003;84:1528-1533.
29. Whitney SL, Wrisley DM, Marchetti GF, Furman JM. The effect of age on vestibular rehabilitation outcomes. *Laryngoscope.* 2002;112:1785-1790.
30. Steffen T, Seney M. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-Item Short-Form Health Survey, and the unified Parkinson disease rating scale in people with parkinsonism. *Phys Ther.* 2008;88:733-746.
31. Blum L, Korner-Bitensky N. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. *Phys Ther.* 2008;88:559-566.
32. Perry J, Garrett M, Gronley JK, Mulroy SJ. Classification of walking handicap in the stroke population. *Stroke.* 1995;26:982-989.
33. Rabadi MH, Blau A. Admission ambulation velocity predicts length of stay and discharge disposition following stroke in an acute rehabilitation hospital. *Neurorehabil Neural Repair.* 2005;19:20-26.
34. Fulk GD, Echternach JL. Test-retest reliability and minimal detectable change of gait speed in individuals undergoing rehabilitation after stroke. *J Neurol Phys Ther.* 2008;32:8-13.
35. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ (Clin Res Ed.).* 2002;324:1417.
36. Peto V, Jenkinson C, Fitzpatrick R. PDQ-39: a review of the development, validation and application of a Parkinson's disease quality of life questionnaire and its associated measures. *J Neurol.* 1998;245 (suppl 1):S10-S14.

37. Deutsch A, Granger CV, Heinemann AW, et al. Poststroke rehabilitation: outcomes and reimbursement of inpatient rehabilitation facilities and subacute rehabilitation programs. *Stroke.* 2006;37:1477-1482.
38. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care.* 1989;27(3)(suppl):S217–S232.
39. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med.* 1993;118:622-629.
40. Stratford PW, Binkley FM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther.* 1996;76:1109-1123.
41. Andrews K, Stewart J. Stroke recovery: he can but does he? *Rheumatol Rehabil.* 1979;18:43-48.
42. Taub E, Uswatte G, Mark VW, Morris DM. The learned nonuse phenomenon: implications for rehabilitation. *Eura Medicophys.* 2006;42: 241-256.
43. Curb JD, Ceria-Ulep CD, Rodriguez BL, et al. Performance-based measures of physical function for high-function populations. *J Am Geriatr Soc.* 2006;54:737-742.
44. Kempen GI, van Heuvelen MJ, van den Brink RH, et al. Factors affecting contrasting results between self-reported and performance-based levels of physical limitation. *Age Ageing.* 1996;25:458-464.
45. Cress ME, Schechtman KB, Mulrow CD, Fiatarone MA, Gerety MB, Buchner DM. Relationship between physical performance and self-perceived physical function. *J Am Geriatr Soc.* 1995;43:93-101.
46. Rozzini R, Frisoni GB, Ferrucci L, Barbisoni P, Bertozzi B, Trabucchi M. The effect of chronic diseases on physical function. Comparison between activities of daily living scales and the Physical Performance Test. *Age Ageing.* 1997;26:281-287.
47. Reuben DB, Valle LA, Hays RD, Siu AL. Measuring physical function in community-dwelling older persons: a comparison of self-administered, interviewer-administered, and performance-based measures. *J Am Geriatr Soc.* 1995;43:17-23.
48. Latham NK, Mehta V, Nguyen AM, et al. Performance-based or self-report measures of physical function: which should be used in clinical trials of hip fracture patients? *Arch Phys Med Rehabil.* 2008;89:2146-2155.
49. Powell LE, Myers AM. The Activities-specific Balance Confidence (ABC) Scale. *J Gerontol A Biol Sci Med Sci.* 1995;50A:M28-M34.
50. Finch E, Brooks D, Stratford PW, Mayo NE. *Physical Rehabilitation Outcome Measures: A Guide to Enhanced Clinical Decision Making*. 2nd ed. Hamilton: BC Decker Inc; 2002.
51. Bayliss MS, Gandek B, Bungay KM, Sugano D, Hsu MA, Ware JE Jr. A questionnaire to assess the generic and disease-specific health outcomes of patients with chronic hepatitis C. *Qual Life Res.* 1998;7:39-55.
52. Quinn L, Gordon J. *Functional Outcomes Documentation for Rehabilitation* St Louis, MO: Elsevier Science; 2003.
53. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. 3rd ed. Upper Saddle River, NJ: Pearson Prentice Hall; 2009.
54. Whyte J. Measuring (treatment-induced) change: conceptual issues. Paper presented at: III STEP Conference on Linking Movement Science and Intervention; July 15-21, 2005; Salt Lake City, UT.
55. Whyte J, Hart T. It's more than a black box; it's a Russian doll: defining rehabilitation treatments. *Am J Phys Med Rehabil.* 2003;82:639-652.
56. Morris DM, Taub E, Mark VW. Constraint-induced movement therapy: characterizing the intervention protocol. *Eura Medicophys.* 2006;42: 257-268.
57. Duncan PW, Jorgensen HS, Wade DT. Outcome measures in acute stroke trials: a systematic review and some recommendations to improve practice. *Stroke.* 2000;31:1429-1438.
58. Whiteneck GG, Charlifue SW, Gerhart KA, Overholser JD, Richardson GN. Quantifying handicap: a new measure of long-term rehabilitation outcomes. *Arch Phys Med Rehabil.* 1992;73:519-526.
59. Wrisley DM, Kumar NA. Functional gait assessment: concurrent, discriminative, and predictive validity in community-dwelling older adults. *Phys Ther.* 2010;90:761-773.
60. Rothstein JM, Echternach JL. *A Primer on Measurement: An Introductory Guide to Measurement Issues*. 1st ed. American Physical Therapy Association; 1993; Alexandria, VA.
61. Beattie P. Measurement of health outcomes in the clinical setting: applications to physiotherapy. *Physiother Theory Pract.* 2001;17: 173-185.
62. Van Genderen F, De Bie R, Helders P, Van Meeteren N. Reliability research: towards a more clinically relevant approach. *Phys Ther Rev.* 2003;8:169-176.
63. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Phys Ther.* 2006;86:735-743.
64. Stevenson TJ. Detecting change in patients with stroke using the Berg Balance Scale. *Aust J Physiother.* 2001;47:29-38.
65. Evans MD, Goldie PA, Hill KD. Systematic and random error in repeated measurements of temporal and distance parameters of gait after stroke. *Arch Phys Med Rehabil.* 1997;78:725-729.
66. Hill KD, Goldie PA, Baker PA, Greenwood KM. Retest reliability of the temporal and distance characteristics of hemiplegic gait using a footswitch system. *Arch Phys Med Rehabil.* 1994;75:577-583.
67. Flansbjer UB, Holmback AM, Downham D, Patten C, Lexell J. Reliability of gait performance tests in men and women with hemiparesis after stroke. *J Rehabil Med.* 2005;37:75-82.
68. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc.* 2002;77:371-383.
69. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol.* 2002;14:109-114.
70. Lang CE, Edwards DF, Birkenmeier RL, Dromerick AW. Estimating minimal clinically important differences of upper-extremity measures early after stroke. *Arch Phys Med Rehabil.* 2008;89:1693-1700.
71. Nilsagard Y, Lundholm C, Gunnarsson LG, Dcnison E. Clinical relevance using timed walk tests and "Timed Up and Go" testing in persons with multiple sclerosis. *Physiother Res Int.* 2007;12:105-114.
72. Beninato M, Portney LG. Applying concepts of responsiveness to patient management in neurologic physical therapy. *J Neurol Phys Ther*. 2011;35(2):75-81.
73. Horak FB, Wrisley DM, Frank J. The Balance Evaluation Systems Test (BESTest) to differentiate balance deficits. *Phys Ther.* 2009;89: 484-498.
74. Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation Systems Test: the mini-BESTest. *J Rehabil Med.* 2010;42:323-331.
75. Sullivan JE, Andrews WA, Lanzino D, Peron A, Potter KA. Outcome measures in neurological clinical practice: Part II. A patient-centered process. *J Neurol Physic Ther*. 2011;35(2):65-74.