Patient- and Clinician-Rated Outcome Measures for Clinical Decision Making in Rehabilitation

Lori A. Michener

Outcome measures can be classified as clinician rated and patient rated. Clinician-rated measures predominantly assess impairments, whereas patient-rated measures, also known as patient-based measures, are designed to evaluate the impact of the injury on a patient’s daily activities, work, and recreation. Currently, there is a greater reliance on clinician-rated impairment measures for clinical decision making, specifically with treatment planning and assessing outcomes of care. To comprehensively evaluate the effect of an injury, patient-rated outcome measures must be used because they allow for the assessment of a patient’s ability to perform daily activities and participate in work and recreation that is affected by an injury. Clinician-rated impairment measures should be used to guide the development of a treatment program, and patient-rated measures should be used for both treatment-program development and assessing treatment outcomes in daily clinical practice. The purposes of this article are to describe patient- and clinician-rated outcome measures and to provide guidance and illustrate the benefits of the use of these measures in clinical decision making and documenting outcomes of care.

Keywords: patient-rated outcome measures (PROMs), orthopedics, sports

The term outcome measures has become common language in the lexicon of health professionals. However, clinicians lack understanding as to the types of outcome measures available, as well as the use and interpretation of these instruments. Outcome measures include measures completed by either the patient or the clinician that are used to guide treatment decision making and evaluate the outcome of an episode of care. Currently, there is greater dependence on the clinician-rated impairment measures than self-report patient-rated outcome measures in daily clinical practice.¹ ² Patient-rated outcomes, also known as patient-based outcomes, are self-reported by the patient and, thus, are designed to assess patient health status from the patient’s perspective, which is ultimately the most important outcome of the care provided. There are a limited number of clinician-rated measures of performance, which evaluate performance of a patient completing functional activities.

Patient-rated outcomes should consistently be part of a comprehensive patient assessment. Unfortunately, studies indicate that only 48% to 50% of physical

The author (lamichen@vcu.edu) is with the Dept of Physical Therapy, Virginia Commonwealth University–Medical College of Virginia Campus, Richmond, VA.
therapists in the United States report use of patient-rated outcomes,\textsuperscript{1,2} and no studies have examined the use of patient-rated outcomes by athletic trainers. Most clinicians (90\%) agree that these measures are important to use, but most do not do so.\textsuperscript{1} Common reasons stated for not using patient-rated outcome measures are unfamiliarity with the measures, belief that they are too time-consuming for patient completion and therapist scoring, and lack of knowledge regarding interpretation and use of the scores.\textsuperscript{1,2} Widespread adoption of patient-rated measures in sport rehabilitation will require education about the different types of outcome tools and how to use them to guide treatment decision making for patients with orthopedic dysfunctions. The purposes of this article are to present an overview of patient- and clinician-rated outcome measures, provide guidance for use and interpretation of outcome measures, and illustrate the benefits of using patient-rated measures with clinician-rated measures to guide treatment decision making.

**Clinician- and Patient-Rated Outcome Measures**

The International Classification of Functioning, Disability, and Health (ICF) was put forth by the World Health Organization as a disablement model to describe the components necessary to comprehensively assess the impact of a disease on a person’s overall well-being.\textsuperscript{3} The ICF defines body functions and structure as physiological functions and anatomical parts of the body and defines impairments as problems in body function or structure. Activity limitations are difficulties in performing a task or action, and participation restrictions are problems experienced when performing activities in the context of a life situation. The ICF categories can be measured with both patient- and clinician-rated measures, as depicted in Table 1.

There are various types of patient-rated outcome measures. Patient-rated outcomes primarily measure activity limitations and participation restrictions and are broadly categorized as either condition-specific or generic measures. Generic measures are not specific to a type of disease or injury; they can be used with all patient diagnoses. In addition, there are generic measures for a variety of health components including pain and general health status. A generic measure of change, the Global Perceived Effect, or Global Rating of Change (GROC),\textsuperscript{4} is a single question that asks patients to rate the level of perceived improvement they have experienced over time. The GROC is appealing, but it has limitations because it requires patient recall of prior health status and it does not query the patient about health-status changes in a variety of activities, making it prone to bias.\textsuperscript{5} Although generic measures are versatile and useful with any patient disease or injury, they can lack specificity of items related to activities and participation for a given patient’s condition. Condition-specific measures, in contrast, are developed to assess a specific disease, specific body part (eg, shoulder, knee), or a body region (eg, lower or upper extremity). A special type of condition-specific measure is the Patient Specific Functional Scale, which asks the patient to list 3 to 5 activity limitations and participation restrictions that are affecting their overall health status.\textsuperscript{6} Patient-specific scales are specific to individual patients and, therefore, can be helpful to assess a small but important number of activity or participation restrictions that are meaningful to the individual patient. Condition-specific measures are recommended for individual patient decision making because they are targeted to the patient’s given disease or injury and important activity limitations and participation restrictions.
Table 1  Patient- and Clinician-Rated Outcome Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>ICF Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinician-Rated Measures</td>
<td>Body structure and function</td>
</tr>
<tr>
<td>Strength</td>
<td>X</td>
</tr>
<tr>
<td>Range of motion</td>
<td></td>
</tr>
<tr>
<td>Joint laxity</td>
<td></td>
</tr>
<tr>
<td>Timed up-and-go&lt;sup&gt;7&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Triple hop for distance&lt;sup&gt;8&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Patient-Rated Measures</td>
<td></td>
</tr>
<tr>
<td>Number of times a shoulder has dislocated</td>
<td>X</td>
</tr>
<tr>
<td>Generic measure: Pain. Rating pain at rest, with tasks, with participation in daily activities: Numeric Pain-Rating Scale.&lt;sup&gt;19&lt;/sup&gt; 0–10 point anchor-based scale; 0 = no pain, 10 = worst pain imaginable</td>
<td>X</td>
</tr>
<tr>
<td>Generic measure: Health status: Medical Outcomes Study Short-Form 36.&lt;sup&gt;22&lt;/sup&gt; 8 dimensions: physical function, role limitations resulting from physical health, vitality, role limitations resulting from emotional health, mental health, bodily pain, social function</td>
<td></td>
</tr>
<tr>
<td>Generic measure: Improvement over time: Global Perceived Effect/Global Rating of Change.&lt;sup&gt;4&lt;/sup&gt; Various scales, 5–15 points; single question asking the patient to rate change in overall status of his or her injury or disease; transitional scale</td>
<td>X</td>
</tr>
<tr>
<td>Condition-specific measures for a specific condition or region of the body</td>
<td></td>
</tr>
<tr>
<td>upper extremity scale: Disabilities of the Arm, Shoulder and Hand&lt;sup&gt;23&lt;/sup&gt;: 30 questions on symptoms and disability rated on a 5-point Likert scale; 0–100 points, 0 = no disability</td>
<td></td>
</tr>
<tr>
<td>lower extremity scale: Lower Extremity Functional Scale&lt;sup&gt;21&lt;/sup&gt;: 20 questions, 5-point Likert scale for difficulty rating; 0–80 points, 80 = full function</td>
<td></td>
</tr>
<tr>
<td>Condition-specific, patient-specific measure: Patient-Specific Functional Scale.&lt;sup&gt;6&lt;/sup&gt; 3 questions that ask the patient to first list 3 activities that are difficult because of his or her condition, and then rate the difficulty level for each activity: 0–10; 0 = unable to perform the activity, 10 = able to perform at same level as before; Total score: 0–30, 30 = full function</td>
<td>X</td>
</tr>
</tbody>
</table>

The measures are described in accordance with the International Classification of Functioning, Disability and Health (ICF) categories and indicate the type of outcome measure. Examples for each are provided.
The generic measure of the GROC can complement the condition-specific measures to judge change over time.

Patient-rated measures are used to comprehensively assess the impact of a disease or injury on a patient across the ICF categories. Clinician-rated measures, which are collected by clinicians, are mostly designed to measure impairments such as limited range of motion or joint laxity. Patient-rated outcome measures are completed by the patient and designed to evaluate the patient’s perspective of the disease or injury, primarily for the ICF domains of activity limitations and participation restrictions. Patient-rated measures must be used to evaluate the outcome of care and guide treatment, because they measure the impact of the disease or injury on the patient’s ability to perform activities and participate in life, work, and sports. Clinician-rated impairment measures should primarily be used to guide treatment decision making with respect to intervention choices. Generally, impairment measures should not be used as a primary outcome measure because they do not take into consideration how difficult it is for patients to use their injured body part during their daily activities, work, or sport.

There are a limited number of clinician-rated performance measures designed to assess activity limitations; the timed up-and-go\(^7\) and the triple hop for distance\(^8\) are 2 examples. Clinician-rated measures of performance may complement the patient-rated measures to evaluate the impact of the disease or injury on the patient’s activities and participation. Emerging evidence indicates that clinician-rated performance measures capture different information about the effect of treatment,\(^9\)–\(^12\) so it may be beneficial to use measures of performance that are both clinician rated and patient rated.

**Interpretation of Scores**

Measures used for treatment decision making and documenting outcomes must have established measurement properties. Specifically, error and responsiveness values are clinically useful because they can be used to interpret a measure’s score. Error values estimate the variability in a measure under 2 conditions: when a measure is used at a single time and when it is used multiple times to assess change over time. The estimate of error with a single administration of a measure is known as the standard error of the measure (SEM); SEM = standard deviation × \((1 – \text{internal consistency coefficient or reliability coefficient})^{1/2}\).\(^{13}\) The minimal detectable change (MDC), also known as the smallest detectable difference, is the error that is associated with 2 administrations of the measure, or change scores; MDC = SEM × \((2)^{1/2}\).\(^{13}\) The SEM and MDC both have a 68% confidence interval (CI). The 90% CIs for the SEM and MDC are calculated by multiplying by the corresponding \(z\)-score of 1.64, producing the 90% confidence of single scores (SEM\(_{90}\)) and that associated with change scores (MDC\(_{90}\)).

Knowing the error values associated with a measure is important because these values are used to interpret the scores from the measure for an individual patient. Moreover, it is essential to understand how to use the error values for both patient- and clinician-rated outcome measures. As an example, a patient who completes the Disabilities of the Arm, Shoulder and Hand (DASH) measure at the initial evaluation has a score of 30/100 (DASH score: 0–100 points, 0 = no disability). Using the SEM\(_{90}\) of 7.4,\(^{14}–^{17}\) the clinician can be 90% confident that the patient’s
score is 30 ± 7.4 points on the DASH. To judge whether a patient is improving or not improving with treatment, measures (eg, DASH) are administered twice during the course of care. Using MDC values can help make sense of how much change in a measure is needed when it is administered 2 times. For example, a patient with a 30/100 DASH score on the initial visit is reassessed with the DASH 2 weeks later and scores a 15/100 on the DASH; this equals a 15-point change on the DASH over 2 weeks. Intuitively, the clinician is left asking what the 15-point change means. The DASH score ranges from 0 to 100 points, with 0 indicating no upper extremity disability. The DASH average MDC_{90} is 10.5 points,\(^{14–17}\) which indicates that a clinician could be 90% confident that true change has occurred with a change score greater than 10.5 points. For the patient example, the 15-point change in the DASH is greater than the DASH’s MDC_{90} of 10.5 points. Therefore, the clinician can be 90% confident that this patient experienced true change and, in this case, improvement.

The MDC is a valuable error statistic to enable the clinician to evaluate a patient’s response to treatment. However, the MDC does not necessarily equate to change that is important to the patient, because it is based on the reliability and variance of the measure. Change that is meaningful to the patient is labeled the minimally clinically important difference (MCID), also known as the minimal important change.\(^4,13,18\) The MCID is the smallest amount of change in a measure that is associated with change that is important or perceived as beneficial to the patient and as such is an index of responsiveness. The MCID value can be used in the same manner as the MDC for clinical decision making. The MCID for the DASH is 10.2 points.\(^{17}\) Using this MCID value, a change of more than 10.2 points on the DASH is associated with meaningful change for the patient. For our patient example, the observed change of a decrease of 15 points on the DASH is considered meaningful change (improvement) in the patient’s level of upper extremity disability as measured by the DASH.

The MCID, SEM, and MDC values help provide meaningful interpretation of outcome measures; however, there are limitations to these values. They are not necessarily stable. The error values and the meaningful change can vary in amount based on such things as the sample of patients used, type of treatment program delivered, interval of change, patient acuity, and baseline outcome score.\(^{13}\) Moreover, athletes have been a small percentage of the patients in studies determining error and responsiveness, which may limit the measures’ use in athletes. In spite of these limitations and variability of the MCID, SEM, and MDC, these values are still quite useful in clinical decision making.

**Use of Patient- and Clinician-Rated Measures to Guide Treatment Decision Making**

Clinician-rated impairment measures are heavily relied on to make decisions regarding patient care. Adding patient-rated measures allows for a comprehensive assessment of the patient’s ability to perform tasks and participate fully in his or her daily life. The 2 types of outcome measures capture different information, so both types have potential to aid decision making. Impairment measures should be used to guide intervention development but not be used as a primary outcome
Impairment measures have a limited ability to determine when a patient is ready to return to work or sport because they only measure impairments such as limited range of motion or strength, but they cannot measure how that impairment affects the patient’s daily activities, work, or sport. Conversely, patient-rated outcome measures do measure activity limitations and participation restrictions and therefore can guide clinician decision making with respect to intervention development and should serve as the primary measure of overall outcome. Outcome of care should focus on patient-rated measures, because they measure what ultimately is most important to the patient. Clinician-rated measures of performance may provide unique information to be used to guide intervention development and serve a limited role as a measure of outcome. The following 2 patient examples exemplify the integrated use of patient- and clinician-rated outcome measures to develop a plan of care, make day-to-day treatment decisions, and determine the end result of treatment.

**Patient Example 1**

The patient is a female soccer player who has had anterior cruciate ligament reconstruction. Over the course of 10 weeks since surgery, she has improved in her level of pain as measured on a numeric pain-rating scale\(^\text{19}\) from a 5/10 to 1/10 (0 = *no pain*) and activity limitations and participation restrictions as per the Lower Extremity Functional Scale (LEFS), with an improvement from a 35/80 to a 50/80 (LEFS range 0–80, 80 = *full function*). She has changed by 4 points on the pain scale, which is a meaningful change with an MCID for the pain scale of 1.2 points.\(^\text{20}\) The LEFS increased by 15 points; this 15-point change is greater than the error in the LEFS and is considered meaningful change, because both the MDC and MCID are 9 points.\(^\text{21}\) Therefore, it is evident that the patient is improving with care, but she still has some loss of activity and participation because her LEFS score of 50/80 (63%) still indicates a deficit.

To determine which interventions should be used to improve activity and participation as measured on the LEFS, the clinician can examine the individual items that are scored high (high level of difficulty) on the LEFS and then assess the impairments that are likely associated with those items. On the LEFS, the soccer player reports difficulty walking and hopping. When she walks across the gym, a lack of full knee extension is observed. Clinician-rated measures reveal limitations of active and passive knee-extension range of motion and quadriceps weakness. Quadriceps weakness and posterior knee-tissue tightness are likely contributing to this soccer player’s deficit in her LEFS score, so treatment directed to correct these impairments should improve the score. Patient-rated measures can facilitate a directed and efficient examination of clinician-rated impairments that are hypothesized to be contributing to activity and participation limitations. After a total of 3 months of treatment aimed at increasing range of motion and improving strength, the LEFS score improved to a 70/80. In addition, the triple hop for distance was added to the battery of tests because it was safe for the patient to perform; her score on the triple hop for distance is 9 less on the affected side than on the unaffected side. This indicates that she is ready to advance her program and start functional activities related to soccer. A combination of clinician-rated performance and patient-rated measures allows for a comprehensive assessment to determine the return to sport.
Patient Example 2

A male lacrosse player with shoulder instability presents with a DASH score of 30/100 (0 = no disability, 0–100 DASH range). The patient reports that his shoulder has subluxed or dislocated twice over the past month since his initial episode of dislocation 1.5 months ago. A treatment regimen directed at improving muscle performance and stabilization was provided for 2 months. After 2 months of this treatment, the patient’s DASH score improved to a 10/100. A change of 20 points on the DASH is greater than the MDC₉₀ of 10.5¹⁴–¹⁷ and the MCID of 10.2 points¹⁷ indicating both a change greater than error in the measure and a meaningful improvement in activity and participation restrictions. The patient reported no episodes of subluxation during the 2-month treatment period. However, the clinician measures of strength via handheld dynamometer indicate no meaningful improvement in strength. There is a meaningful improvement in the DASH to a 10/100 and a reduction in the incidence of subluxation. Overall, these findings suggest that although the patient still has a strength deficit, the deficit appears not to be large enough to have an impact on his ability to perform his daily activities and participate in work and sport. He notes he has been practicing with his lacrosse stick and throwing the ball below shoulder level only. Thus, this patient is likely ready for progression in his lacrosse activities and functional retraining to return to playing lacrosse.

Summary

Patient-rated measures must be used to comprehensively assess the impact of disease or injury on a patient’s activity limitations and participation restrictions. Patient-rated measures should be completed across the continuum of care to guide treatment decision making and assess the outcome of care. Using error and meaningful change values allows for easy interpretation of the scores obtained from the patient-rated outcome measures. To guide intervention choices and planning, patient-rated along with clinician-rated measures should be used. The primary measure of treatment outcome should be patient-rated outcome measures, not clinician-rated impairment measures. There is emerging evidence indicating that clinician-rated performance measures may provide additional unique information with respect to outcome of care. In orthopedics, patient-rated outcomes are often superseded by clinician-rated measures of impairment in evaluating and managing patients. We suggest that patient-rated measures for clinical decision making and documenting outcomes of care in sport rehabilitation need to be incorporated as standard practice. Are you ready?

Acknowledgments

This article is part of a special issue of JSR on clinical outcomes assessment.

References


