

# Selecting Patient-Based Outcome Measures

Alison R. Snyder, PhD, ATC and Tamara C. Valovich McLeod, PhD, ATC •  
A.T. Still University

**C**LINICAL OUTCOME ASSESSMENT involves the use of clinician-based and patient-based measures. Athletic trainers are most familiar with clinician-based measures, because they frequently relate to impairments that are recorded for the “objective” section of SOAP-format documentation (e.g., range of motion and strength). Patient-based measures are typically self-report questionnaires that ask questions regarding the patient’s perception of his or her condition, injury, and/or overall health status. Ideally, clinician-based and patient-based outcome measures should be included in all initial and follow-up patient evaluations; however, the large number of available instruments may overwhelm clinicians who have some degree of apprehension about collecting patient-based outcome measures.

Patient-based outcome measures are classified as either generic or specific. Generic measures are suitable for a wide variety of patient populations and injuries; they are broad in scope and typically focus on health-related quality of life.<sup>1,2</sup> In contrast, specific outcome measures focus on a particular region, disease/condition, or health dimension.<sup>1,2</sup> Region-specific scales address the effects of injury/illness on a specific body region (e.g., shoulder), whereas disease/condition-specific measures are useful for evaluation of medical conditions (e.g., shoulder instability). Dimension-specific instruments are focused on one aspect of an individual’s health, such as pain. Optimal patient care requires evaluation of both generic and specific outcomes; however, there are far fewer generic measures than specific measures. For example, there are over 25 patient-based shoulder-specific outcome measures.<sup>3</sup> It is important for the practicing clinician to determine

which of these instruments is most appropriate for his or her clinical practice, or most appropriate for individual patients, which is challenging. The purpose of this report is to present eight criteria deserving consideration when selecting a patient-based outcome measure (Table 1).

## Instrument Development

The initial concern regarding any patient-based outcome measure is the procedure used for instrument development. Development of patient-based measures is complex, involving item generation and initial item reduction, field testing and final item reduction, and establishment of scale measurement properties.<sup>4,5</sup> A clinician should determine whether a systematic process was used to develop this instrument.<sup>5</sup> For example, the International Knee Documentation Committee Subjective Knee Form (IKDC Subjective Knee Form) was developed through a comprehensive process that began with the identification of the purpose of the instrument and the constructs to be measured.<sup>6</sup> Questions for the initial version of the instrument, a total of 27, were developed by the committee through review of other knee instruments, with the goal of highlighting symptoms and functional capabilities related to activities of daily living and sports. Following question generation, a series of procedures were performed, including field testing of preliminary questions, determination of measurement properties of the initial questions, identification of final instrument questions, and field testing of the final instrument.<sup>6</sup> As a result of the systematic development process, the IKDC Subjective Knee Form is a reliable and valid instrument for

**TABLE 1. EIGHT CRITERIA FOR SELECTING PATIENT-BASED OUTCOME MEASURES**

1. Instrument development
2. Appropriateness
3. Reliability
4. Validity
5. Responsiveness
6. Interpretability
7. Acceptability
8. Feasibility

quantification of symptoms and function for a wide variety of people suffering from injuries and conditions affecting the knee.<sup>6</sup> The IKDC is an example of a well-developed patient-based outcome measure practical for use by athletic trainers.

### Appropriateness

Appropriateness of a scale refers to how closely the measure relates to a specific clinical question.<sup>7</sup> A clinician should identify a particular purpose for the use of the scale (e.g., rehabilitation goal), which should relate to the injury, disease/condition, or dimension of interest. A clinician should determine whether the content of the instrument is appropriate for the clinical question, condition, or patient being treated, and whether the instrument is appropriate for a variety of patients seen in clinical practice. For example, the Anterior Knee Pain Scale (AKPS)<sup>8</sup> is an appropriate evaluation tool for a patient with anterior knee pain related to patellofemoral pain syndrome and dislocation or subluxation of the patella; however, the AKPS is limited to questions related to pain and function of the knee. In contrast, the Lower Extremity Functional Scale (LEFS)<sup>9</sup> is broader in scope, and is useful for a variety of patients and lower extremity conditions. The LEFS is appropriate for patients with hip osteoarthritis, ankle fractures, patellofemoral knee pain, and muscle strains. In contrast, the AKPS is not appropriate for people with hip osteoarthritis or ankle fractures.

Clinicians should not select an instrument on the basis of title alone, because the title cannot precisely identify the construct that is measured by an instrument.<sup>7</sup> In terms of physical function, one instrument may ask questions specific to the amount of assistance a person needs to complete particular tasks, whereas another scale may ask questions regarding the degree of difficulty the person has in completing tasks. The representations of physical capabilities derived from these instruments may not be the same. A review of the instruments, beyond title alone, will assure that an instrument is appropriate for its intended use.

### Reliability

Reliability is the degree to which an instrument measures the true scale score, as opposed to random error.<sup>3,7,10,11</sup> Clinicians should determine whether the instrument produces results that are reproducible and internally consistent. Reproducibility or “stability” can be measured through administration of the instrument on two separate occasions, during an interval in which patients’ conditions are not expected to change.<sup>3,7,10,11</sup> Typically, reproducibility is represented by the Interclass Correlation Coefficient (ICC), with values above 0.7 considered acceptable.<sup>7</sup> Internal consistency relates to the homogeneity of questions in the same domain and their ability to measure the same construct.<sup>3,7</sup> Internal consistency is frequently evaluated with Cronbach’s alpha, which is generally deemed acceptable at values of 0.7-0.9.<sup>7</sup>

### Validity

Validity refers to the ability of an instrument to measure what it intends to measure.<sup>3,7,10</sup> Several types of validity are established for instruments, and clinicians should determine whether a particular instrument demonstrates an acceptable range of validity for a target population<sup>7</sup> (Table 2). Although an instrument may be validated for a particular population or injury, validity is not directly transferable to other populations or injuries. The Disabilities of the Arm, Shoulder, and Hand (DASH) scale is well validated for patients ranging in age from 40-58 years old with upper extremity complaints,<sup>3</sup> but to our knowledge, there are no studies validating its use for an adolescent athletic population.