

# Techniques for Revalidating the Reliability and Validity of Patient Reported Outcome Measures

Daniel J. Simmons, Rutgers, the State University of New Jersey, New Brunswick, NJ, USA



Daniel J. Simmons

## KEY POINTS . . .

PROs are generally validated when they are made, but may need to have reliability and validity revalidated when used for a new population or in a new way.

Reliability measures the degree to which a score on a PRO reflects actual differences.

Validity demonstrates that a test can be interpreted as measuring what it claims to measure.



*The following article is the fifth in a series highlighting local student chapter activities and research talents. In this piece, we review evaluating patient-reported outcomes accuracy.*

## What are PROs and When Do They Need to be Revalidated?

Patient-reported outcomes (PROs) are important measures in health care and health outcomes research. Fundamentally, PROs represent any outcome that a patient directly reports to researchers [1,2]. PROs differ from clinical endpoints in that they take into account the impact of a condition or intervention on the life of a patient from the point of view of the patient. Commonly measured PROs include symptoms, quality of life, adherence to a treatment, satisfaction, and patient functioning [1,3]. Because they can evaluate patient burden, PROs are especially useful in studying diseases such as arthritis, where there are limited clinical measures that accurately describe the impact of the disease, and in diseases such as cancer where treatments can be difficult to tolerate [1,2].

Since the definition of PROs is broad and can encompass many different measures, it is essential to utilize a method of standardization to ensure that results of PRO measures can be interpreted. When a new PRO questionnaire or measure is created, it generally undergoes a strong psychometric validation process to ensure that it is valid and reliable [1,4]. There are, however, scenarios when a previously validated PRO measure should be revalidated, including changes to the content of the measure, the way the measure is administered (e.g., via the internet vs. in person), the language of administration, or significant differences in the patient population [4].

## Reliability

Assessing the reliability of a PRO is important when revalidating the measure. In classical test theory, the concept of reliability can be summarized as the degree to which the score on an assessment measures actual differences in a trait and not measurement error or random chance [5]. A test with a high degree of reliability

should have consistent findings each time it is administered to the same subject [1]. There are multiple methods used to assess reliability, such as measuring internal consistency reliability and test-retest reliability [1,5].

Internal consistency reliability is commonly assessed for measures that have multiple questions or items and yield an overall score. Since calculating internal consistency reliability only requires one administration of the test, it is especially useful for measuring reliability during revalidation [5]. The theory behind internal consistency reliability is that in a reliable multi-item test of the same construct, all of the items should have strong associations. The most common method to test for internal consistency reliability is Cronbach's  $\alpha$  [5]. There are multiple formulas that can be used to measure Cronbach's  $\alpha$  that take into account the sum of covariance between items, the number of items, and item variances [1-5]. For a large study, the best way to calculate Cronbach's  $\alpha$  is to use a statistics program such as SPSS, STATA, SAS, or R. These programs allow a researcher to simply enter the question number and response or coded response for each subject. In SPSS, STATA and R, there are options for Cronbach's  $\alpha$ , and in SAS it is under "Cronbach Coefficient Alpha for Raw Variables." [5] The coefficient produced from this computation will range from 0 to 1 with values  $\geq 0.7$  being generally recognized as acceptable reliability [1,6].

Calculating test-retest reliability is another key method of evaluating reliability. If a test is reliable, there should be a strong correlation between scores for each subject when the test is administered multiple times [5]. The time between administrations should be limited enough so that there is minimal variation due to the underlying condition or intervention, but not so limited that the subject recalls their exact answers [1-5]. The standard measure of test-retest reliability recommended in the Patient-Centered Outcome Research Institute (PCORI) guidelines is the intra-class correlation coefficient (ICC) [6]. Most statistical software including SPSS, STATA, SAS, and R allow for direct calculation of the >

ICC. The common threshold for acceptable test-retest reliability is an ICC  $\geq 0.7$  [6]. Test-retest reliability may be a bit more challenging to use for revalidation because it requires multiple administrations of the test. It is also not ideal for constructs that change frequently such as mood [5].

## Validity

In addition to reliability, validity is important in revalidating PROs. Validity can be understood as the extent to which a score on an assessment can be interpreted as measuring the construct it claims to measure [1-5]. Validity is a continuum and does not contain a specific score that delineates a valid test from an invalid test [5]. Additionally, there are multiple types of validity including face validity, content validity, criterion validity, and construct validity [1-5].

Construct validity is often investigated when revalidating a PRO. It is the extent to which a measure behaves compared to an expected hypothesis [1]. If a measure has strong construct validity, the scores on the measure should correlate with other validated measures used on the same subjects [5]. This means that if a test is valid, its score should correspond with other clinical or PRO measures hypothesized to be related. Many techniques to evaluate construct validity can be used including correlations, regression, and factor analysis [1]. To test the construct validity, statistical software can be used to measure correlations between scores for the PRO and previously validated measures. When conducting analysis with SPSS, STATA, SAS, or R, researchers should look for statistically significant correlations that have a P value  $\leq 0.05$  [1].

## Conclusion

PROs have the ability to provide important endpoints in research studies. When modifying a PRO or using it for a new purpose or population, it is important to revalidate the measure to ensure that it is consistent and measures what it claims to measure. Validity and reliability are important concepts to understand and evaluate when beginning to revalidate the measure. Once revalidated, the PRO has the potential to provide more meaningful results because it can be viewed as valid and reliable.


## Acknowledgements

Editing assistance provided by Christina A. Simmons, PhD, Munroe-Meyer Institute, University of Nebraska Medical Center, Omaha, NE, USA.

## References

- [1] Frost MH, Reeve BB, et al. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? Value Health doi:10.1111/j.1524-4733.2007.00272.x. [2] Patrick DL, Burke LB, Powers JH, et al. Patient-reported outcomes to support medical product labeling claims: FDA perspective. Value Health 2007 doi:10.1111/j.1524-4733.2007.00275.x. [3] Wiklund I. Assessment of patient-reported outcomes in clinical trials: the example of health-related quality of life. Fundam Clin Pharmacol 2004;18:351-63. [4] Snyder CF, Watson ME, Jackson JD, et al. Patient-reported outcome instrument selection: Designing a measurement strategy. Value Health 2007;10(Suppl.):S76-S85. [5] Furr MR, Bacharach VR. Psychometrics: An introduction. Los Angeles: Sage Publications; October 19, 2007. [6] Acaster S, Cimms T, Lloyd A. The design and selection of patient reported outcomes measures (PROMS) for use in patient-centered outcomes research. Oxford Outcomes/PCORI; 2012:1-82. ■

< ADVERTISEMENT >



**RTI (h)(s)<sup>TM</sup>**  
**Health Solutions**

**Building our greatest asset. Our people.**

**Browse our open positions and apply online at**  
[rtihs.org/ISPORvos](http://rtihs.org/ISPORvos)

We're always looking for talented individuals to join our team of academically trained researchers and experienced professionals. We seek people of intellect, creativity, and commitment who understand the vital connection between focused research and real-world results, and who realize that furthering a career in applied science and study can provide a significant benefit to society.

**Research Triangle Park | Ann Arbor | Barcelona | Ljungskile | Manchester | Waltham**