

White Paper | PRO Consulting

Conceptualizing Modifications and Identifying Validation Procedures for PRO Instruments

Written by:
Alan Shields, Ph.D.
Chad Gwaltney, Ph.D.
Jean Paty, Ph.D.
Saul Shiffman, Ph.D.

Date:
August 22, 2006

© 2006 PRO Consulting. All rights reserved.



Introduction

According to the Food and Drug Administration (FDA) Draft Guidance on Patient Reported Outcomes (PRO) measures, additional validation evidence may be required when PRO measures are modified for use in a clinical trial.¹ The Draft Guidance states, “The extent of additional [PRO instrument] validation recommended depends on the type of modification made” (lines 582–583). Thus, clinical researchers must determine the degree of (re)validation required when modifications are made to a PRO measure.

These are the primary goals of this paper:

- 1 Identify the significance or magnitude of modifications to PRO measures.
- 2 Identify the procedures that may be necessary to validate those changes.

A PRO Instrument Validation Hierarchy

Consistent with the science-based and flexible tone that characterizes the FDA’s PRO Guidance, we developed a PRO instrument Validation Hierarchy that matches the “level” of instrument modification to a specified validation procedure (Box 1). When a researcher decides that a PRO instrument modification is necessary, our Validation Hierarchy prompts him or her to categorize that change among four levels of PRO instrument modifications. This categorical decision, in turn, suggests a particular validation procedure to validate that change.

While not exhaustive (all possible instrument modifications and/or validation procedures could not be accounted for within the model), this Validation Hierarchy can be a useful starting point for researchers using PRO instruments in their existing and/or future clinical trials. It can guide how researchers begin to conceptualize their instrument modifications and inform them of the procedures they may need to employ to produce the validation evidence the FDA recommends for instrument modifications.

Box 1: PRO Instrument Validation Hierarchy

Level of Modification	Justification	Examples	Level of Evidence
Small	The modification can be adequately justified based on logic and/or existing empirical evidence. No change in content or meaning.	<ol style="list-style-type: none"> 1) Changing font, font size, color of paper or color of ink (so long as result is readable) 2) Embedding in a battery of instruments 3) Changing from a "circled" response option (paper) to a "tapped" one (electronic) 	None
Medium	The modification can be adequately justified based on existing empirical evidence. No change in content or meaning.	<ol style="list-style-type: none"> 1) Language or cultural translation 2) Minor changes in mode of administration (e.g., from paper to electronic diary) 	Cognitive debriefing
Large	The modification cannot be justified as neutral or based on existing empirical evidence. May change content or meaning.	<ol style="list-style-type: none"> 1) Changes in item wording that alter interpretability 2) Large change in mode of administration (e.g., from paper to interactive voice response) 	Equivalence testing
Substantial	The modification cannot be justified as neutral or based on existing empirical evidence. Changed content or meaning.	<ol style="list-style-type: none"> 1) Substantial changes in item response options 2) Substantial changes in item wording 3) Computer-adaptive testing 	Psychometric revalidation study

Conceptualizing PRO Instrument Modifications

A large body of research informs decisions regarding the significance of a modification to a PRO measure. Differences in the data derived from a modified PRO instrument (relative to an unmodified version) are driven by the extent to which the modifications alter how the respondent understands or interprets individual test questions.² This suggests the following:

- ④ Changing the font, color or size of the paper used for pen-and-paper PRO instruments would not theoretically alter how respondents interpret a question as long as the change did not affect readability.
- ④ Changing the content or meaning of items and/or response options will alter how respondents interpret a question.³

Thus, to categorize the level of change made to a PRO instrument (column 1, Box 1), researchers should ask whether that change alters the content and/or meaning (i.e., the interpretability) of the original instrument. If the modification does not alter the content and/or meaning of the original instrument, it may be labeled a “small” or “medium” change. If, however, the modification does alter the content and/or meaning of the original instrument, then labeling that modification as “large” or “substantial” may be more appropriate. Notably, in this case, the procedure necessary to validate the identified modification depends on its “level of modification.”

Identifying Validation Procedures

Column 4 of Box 1 reflects four validation procedure domains ranging from less to more intensive:

- ④ No revalidation necessary (use existing evidence)
- ④ Cognitive debriefing
- ④ Equivalence testing
- ④ Psychometric validation and revalidation studies

After a researcher decides that a PRO instrument modification is necessary and has categorized that change among the four levels of PRO instrument modifications (see above), the Validation Hierarchy matches a particular validation procedure to that change.

In many cases it will be useful to survey the literature to determine if prior studies have examined the effect of a particular modification. For example, a recent meta-analysis found that porting a PRO instrument from paper to electronic administration on a PC or palmtop platform yields a psychometrically equivalent instrument, thus mitigating the need to test this (i.e., revalidate) every time a new instrument is migrated from paper to electronic administration.⁴ The practical advantage of this is that researchers porting their paper PRO tools to electronic text platforms — regardless of the tool — can be confident that the migration will produce equivalent results and that they will not have to invest resources into more expensive and time-consuming revalidation procedures.

Conclusions

The FDA recognizes that different degrees of modifications of PRO instruments require different degrees of revalidation. This flexible point of view is logical and consistent with decades of psychometric and applied research showing that small changes have little effect on an instrument, whereas larger ones can undermine its validity. Corresponding to the range of modifications, the PRO Guidance also recognizes a variety of validation procedures, ranging from less to more resource-intensive. For example, validation evidence could be obtained via existing studies, cognitive debriefing, equivalence testing and/or full psychometric validation studies. The Validation Hierarchy presented within this paper can serve as a useful starting point for researchers who are concerned about what evidence may be necessary to document the validity of their modified PRO instruments and how they might go about acquiring and documenting that evidence.

References

- ¹ Federal Register, Vol. 71, No. 23; Feb. 3, 2006.
- ² Schwarz, N. (1999), "Self-reports. How the questions shape the answers," *American Psychologist*, 54, 93–105.
- ³ Winkielman, P., Knauper, B., and Schwarz, N. (1998), "Looking back at anger: Reference periods change the interpretation of (emotion) frequency questions," *Journal of Personality and Social Psychology*, 75, 719–728.
- ⁴ PRO Consulting (unpublished white paper, 2006), "Equivalence of Electronic and Paper-and-Pencil Administration of Patient Reported Outcomes: A Meta-Analysis," <http://www.patientreported.com>.