VALUE IN HEALTH 14 (2011) 967–977

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/jval

ISPOR TASK FORCE REPORTS

# Content Validity—Establishing and Reporting the Evidence in Newly Developed Patient-Reported Outcomes (PRO) Instruments for Medical Product Evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 1—Eliciting Concepts for a New PRO Instrument

*Donald L. Patrick, PhD, MSPH[1,*], Laurie B. Burke, RPh, MPH[2], Chad J. Gwaltney, PhD[3], Nancy Kline Leidy, PhD[4], Mona L. Martin, RN, MPA[5], Elizabeth Molsen, RN[6], Lena Ring, PhD[7]*

[1]*Department of Health Services, University of Washington, Seattle, WA, USA;* [2]*Office of New Drugs, Center for Drug Evaluation Research, Food and Drug Administration, Silver Spring, MD, USA;* [3]*Department of Community Health, Brown University, Providence, RI, USA, and PRO Consulting, Pittsburgh, PA, USA;* [4]*United BioSource Corporation, Bethesda, MD, USA;* [5]*Health Research Associates, Inc., Seattle, WA, USA ;* [6]*International Society for Pharmacoeconomics and Outcomes Research, Lawrenceville, NJ, USA;* [7]*Health Economics & Outcomes Research Division, AstraZeneca, Södertälje, Sweden, and Pharmaceutical Outcomes Research, Department of Pharmacy, Uppsala University, Uppsala, Sweden*

ABSTRACT

The importance of content validity in developing patient reported outcomes (PRO) instruments is stressed by both the US Food and Drug Administration and the European Medicines Agency. Content validity is the extent to which an instrument measures the important aspects of concepts that developers or users purport it to assess. A PRO instrument measures the concepts most significant and relevant to a patient's condition and its treatment. For PRO instruments, items and domains as reflected in the scores of an instrument should be important to the target population and comprehensive with respect to patient concerns. Documentation of target population input in item generation, as well as evaluation of patient understanding through cognitive interviewing, can provide the evidence for content validity. Developing content for, and assessing respondent understanding of, newly developed PRO instruments for medical product evaluation will be discussed in this two-part ISPOR PRO Good Research Practices Task Force Report. Topics include the methods for generating items, documenting item development, coding of qualitative data from item generation, cognitive interviewing, and tracking item development through the various stages of research and preparing this tracking for submission to regulatory agencies. Part 1 covers elicitation of key concepts using qualitative focus groups and/or interviews to inform content and structure of a new PRO instrument. Part 2 covers the instrument development process, the assessment of patient understanding of the draft instrument using cognitive interviews and steps for instrument revision. The two parts are meant to be read together. They are intended to offer suggestions for good practices in planning, executing, and documenting qualitative studies that are used to support the content validity of PRO instruments to be used in medical product evaluation.

***Keywords:*** content validity, European Medicines Agency, Food and Drug Administration, patient reported outcomes, quality of life.



## Background to the Task Force

In March 2009, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Board of Directors approved the formation of the Patient Reported Outcomes (PRO) Content Validity Good Research Practices Task Force to develop a good research practices report to address methods for ensuring and documenting the content validity of newly developed PRO instruments to support medical product indications and labeling claims. This task force report extends the work of a previously published ISPOR PRO task force report on the use of existing or modified PRO instruments [1], which did not address how to establish and document content validity; that is, the specific methodologic practices involved in designing studies to gather evidence of content validity and the methods for evaluating and documenting content validity.

Researchers experienced in psychometrics and PRO instrument development working in academia, government, research organizations, and industry from North America and Europe were invited to join the task force leadership group. The task force met bimonthly to develop the topics, outline, and prepare the first draft report. Due to the large volume of information, the task force report was split into two parts. Part 1 covers elicitation of key concepts using qualitative focus groups and/or interviews to inform content and structure of a new PRO instrument. Part 2 [2] covers the instrument development process, the assessment of patient understanding of the draft instrument using cognitive interviews and steps for instrument revision.

Authors listed in alphabetical order after lead author.

* *Address correspondence to:* Donald L. Patrick, University of Washington—Health Services, PO Box 359455 SeaQoL Group, Seattle, WA 98195-9455, USA.

E-mail: donald@u.washington.edu.

1098-3015/$36.00 – see front matter Copyright © 2011, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

doi:10.1016/j.jval.2011.06.014

The task force authors presented their work to date at the ISPOR 15th Annual International Meeting in May 2010 in Orlando, Florida. During July 2010 the draft report papers (Part 1 and Part 2) were sent for review to the nearly 400 ISPOR PRO Review Group members. The task force received many comments that were considered and addressed as appropriate. The task force authors presented their revised draft report for final verbal comments at the ISPOR 16th Annual International Meeting in Baltimore, Maryland, during May 2011. The revised report was sent for a final review to all ISPOR members during June 2011.

Collectively, the task force received 41 written reviews by 52 ISPOR members submitted individually or representing an organization. All written comments are published at the ISPOR Web site. A list of those members who commented is available. For these comments, please go to the 'Evaluating and Documenting Content Validity for PRO Instruments' link at the ISPOR Good Outcomes Research Practices index under the Patient Reported Outcomes heading at: http://www.ispor.org/workpaper/practices_index.asp or via the purple Research Tools menu at the top of the ISPOR homepage (www.ispor.org). All comments, many of which were substantive and constructive, were considered. Once consensus was reached by all authors, the final report was submitted to *Value in Health* in July 2011.

## Introduction

According to the US Food & Drug Administration (FDA) [3], a PRO is "any report of the status of a patient's health condition that comes directly from the patient without interpretation of the patient's response by a clinician or anyone else." It can be measured in absolute terms (e.g., severity of a sign, symptom, or state of a disease) or as a change from a previous measure [3].

The European Medicines Agency's (EMA) Reflection Paper on the Regulatory Guidance for the Use of Health Related Quality Of Life (HRQL) Measures in the Evaluation of Medicinal Products [4] defines a PRO similarly as "any outcome directly evaluated by the patient and based on patient's perception of a disease and its treatment(s)" [4]. EMA uses PRO as an umbrella term encompassing both single and multidimension domains; that is, measures of symptoms, health status, and satisfaction with treatment [5]. HRQL is one of these multidimensional assessments under the PRO heading, broadly defined as a patient's subjective perception of the effects of the disease and treatment(s) on daily life; well-being; and psychological, physical, and social functioning. In the drug approval context, HRQL is considered a specific type of PRO [4].

Because the term PRO is often used interchangeably to refer to a PRO concept, questionnaire, instrument, score, or claim, it is useful to define these terms. PRO is the general reference to the concept (outcome) of interest. The PRO field is the general area of study. Elements of the field include PRO research; for example, burden of illness studies, qualitative theory-development studies, clinical trials, instrument development research, and PRO instrument development. Instrument development comprises the qualitative and quantitative studies that identify and measure outcomes reported by patients themselves.

A PRO instrument (i.e., a questionnaire plus the information and documentation that support its use) is a means to collect data about a PRO concept. A PRO instrument extends patient outcome assessment beyond survival, traditional clinical efficacy, and adverse effects. It assesses the concepts most relevant and important to a patient's condition and treatment. A PRO measure refers to a specific questionnaire used to collect data that produces a score representing the PRO concept of interest.

In medical product development, PRO instruments may be used in clinical trials to capture and quantify treatment benefit or risk [3,6]. This information potentially may be used to support a claim in medical product labeling or advertising. Within this context, it is useful to distinguish the PRO concept, claim, instrument, and score [6]. For example, pain intensity is a PRO (i.e., the concept), whereas a decrease in pain intensity might be a PRO claim based on a prespecified endpoint in a clinical trial. A 10-centimeter visual analog scale that assesses pain intensity—including the anchors, instructions, and recall period—is a PRO instrument. Finally, the value a subject assigns to pain intensity on the visual analog scale is a PRO score.

PRO instruments are designed to capture concepts related to the health experiences of individuals—how patients feel or function in relationship to their disease, condition, or treatment. Thus, the instruments must possess content validity. In the FDA Guidance on PRO measurement, content validity is defined by the empiric evidence that demonstrates the items and domains of an instrument are appropriate and comprehensive relative to its intended measurement concept, population, and use [3]. In practical terms, content validity is determined by documenting that the structure and content (items) capture the connection between the intended measurement concept and the way patients from the target population understand and discuss that concept. A full description of these methods and their results provides evidence that scores produced by the instrument represent the intended concept; that is, are content valid.

Qualitative data are essential for establishing the content validity of a PRO instrument. Quantitative data, including factor analysis, Rasch analyses, or item response theory analyses may be supportive, but such data are insufficient on their own to document content validity of the measure in the context of medical product development. Content validity must be based on direct input from an adequate sample of patients from the targeted clinical study population. Involving a diverse sample helps ensure that the final instrument measures the intended concept despite important variations in demographic and clinical characteristics and experiences within the target population.

Part 2 of this report describes the second phase of establishing and reporting evidence of content validity for a new PRO instrument—its development and the methods for gathering evidence that persons in the target population understand the instrument's structure and content [2].

## Good Practices in Eliciting Concepts for a New Pro Instrument

Table 1 lists five steps to elicit concepts for establishing and documenting content validity of a new PRO instrument, consistent with the wheel and spokes diagram presented by the FDA [3]. These five steps represent the initial stages of instrument development. The development process in general is an iterative, rather than linear process, often requiring researchers to revisit previous steps to ensure adequate and accurate information related to instrument content or structure and to fully document content validity relative to the context of use. Each of these steps is described below.

### *Good practice 1: Determine the context of use*

The development of an instrument, whether simple or complex, starts with the identification of the concept and the target medical product labeling claims, so those targets can be considered throughout the instrument development process. The purpose of Step 1 in Table 1 is to ensure that the context of use in medical product labeling is clearly defined, and the approach for concept measurement is appropriate for the intended context. This includes an understanding of the disease or condition in the target population, development of an endpoint model for the context of use, considerations related to specific aspects of the target population, the possible range of instrument content and structure, the theoretical and qualitative methodologic approach, and the development of a hypothesized conceptual framework. Each of these are discussed below.