Study Endpoint Considerations: Final PRO Guidance and Beyond

Laurie Burke
Associate Director for Study Endpoints and Labeling
OND/CDER/FDA

Presented at:

FIRST ANNUAL

PATIENT-REPORTED OUTCOMES (PRO) CONSORTIUM WORKSHOP

March 23, 2010 – Bethesda, MD

Co-sponsored by

Overview

- What is endpoint measure qualification?
- Why qualify endpoint measures?
- Standards of qualification evidence: PRO guidance
- Qualification submissions: Current FDA review experience
- Beyond the PRO experience
  - Non-PRO measures (e.g., caregiverROs, ClinROs)
  - Other purposes beyond labeling claims
    - Prospective safety assessment
    - Other

PRO Qualification: What?

- Regulatory conclusion that within the stated context of use, the results of measurement can be relied upon to have a stated interpretation and utility—”fit for purpose”
- Data produced by the PRO measure (i.e., instrument) can be interpreted as clinically meaningful and can be used as a primary or key secondary endpoint to support a claim in labeling
- Meets the standards set forth in the PRO guidance

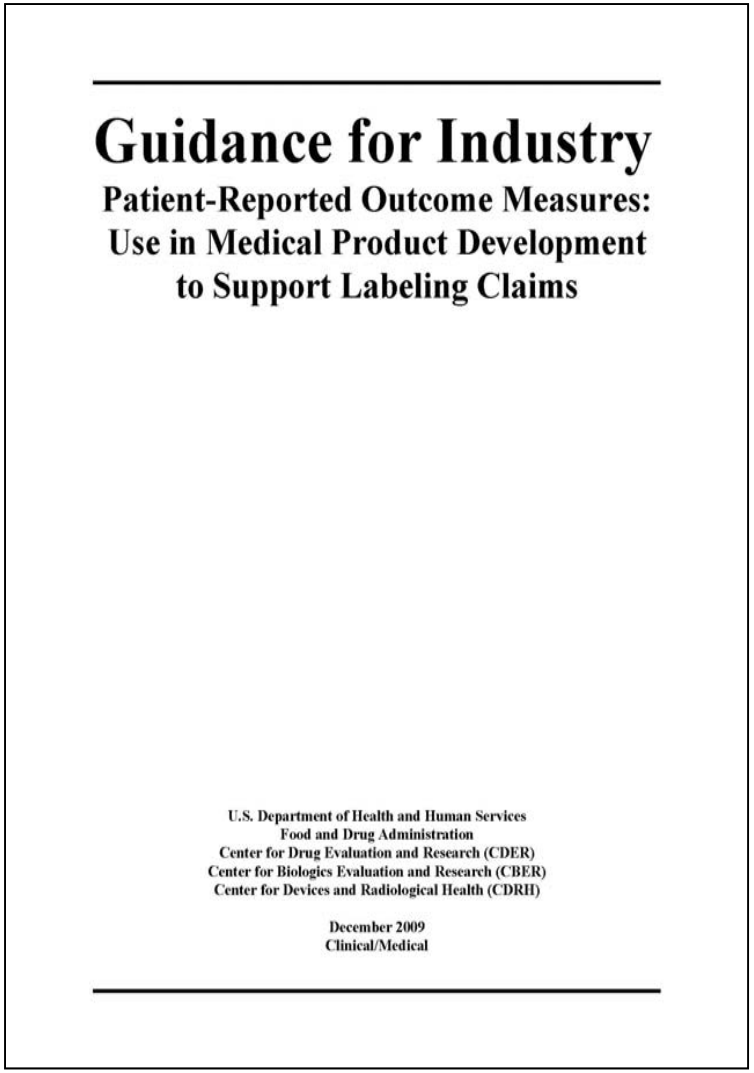
PRO Qualification: Why?

- PRO measures provide valuable information to support claims for product labeling
- PRO measures provide efficiency in clinical study design
- Lack of well-established PRO measures for many diseases where symptoms are important indicators of clinical benefit

qualification

- Efficiency for industry/FDA time and resources
- Availability of PRO measures in the public domain
- A more transparent advisory process
- Heightened awareness of good measurement principles
- Better information for patients and other decision-makers

PRO Qualification: Standard of evidence



Guidance for Industry

Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Center for Devices and Radiological Health (CDRH)

December 2009
Clinical/Medical

Feb 2006: Draft

Dec 2009: Final

http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf

PRO Guidance

- Explains how FDA evaluates PRO measures for their usefulness in measuring & characterizing treatment benefit as perceived by the patient.
- Explains how FDA reviews and interprets evidence that a PRO instrument measures the concept represented by a treatment benefit claim.

When Is a PRO Instrument "Fit for Purpose?"

- **Content validity is established for the intended purpose**
  - **Measurement concept matches targeted claim**
  - **Item content development includes target population input (qualitative research)**
  - **Item content captures the intended concept in the intended treatment population**
  - **Measurement concept conforms with the proposed clinical trial objectives**
- **Construct validity, reliability and ability to detect change are adequate to support interpretation of clinical trial results**
- **User manual provides adequate instructions for use and scoring**

Scoping document: Start with the end in mind

- Identify the measurement goals: Concept(s)
- Identify the target population
  - Disease severity
  - Cultural/language groups
- Identify the role of the measure in context of other clinical trial endpoints
  - Endpoint model
  - Clinical trial design
- Hypothesize preliminary conceptual framework of the instrument

Measurement Properties: Content Validity

- The extent to which the content of an instrument represents most important aspects of a given concept based on empiric evidence generated from patients from the target population
  - Response options capture full range of severity
  - Recall period is appropriate for the concept and population
  - Signs and symptoms are distinguishable from impacts
  - Mode of administration is appropriate
  - Patients understand item content and instructions for use

Content Validity: Based on Qualitative and Quantitative Research

- Literature review and expert input on concept, disease, existing instruments
- Guided by qualitative research protocols and analysis plan
- Listen to patients (focus group testing or open-ended patient interviews)
- Testing of draft instruments using cognitive patient interviews
- Statistical exploration of qualitative data
  - Text analysis software
  - Factor analysis
  - Item response analysis

Content Validity: Documentation

- Hypothesized conceptual framework of the instrument
- Summary of results of concept elicitation and cognitive patient interviews (focus group or individual)
  - Characteristics/demographics of study population
  - Evidence of saturation for concept elicitation
  - Transcripts available upon request
- Origin and derivation of:
  - Item content
  - Response options
  - Recall period
  - Scoring
- Chronology of events
  - Item generation
  - Item modification
  - Finalization of item content
- Final conceptual framework of the instrument

Measurement Properties: Construct Validity

- Evidence that relationships among items, domains, and concepts conform to *a priori* hypotheses concerning logical relationships that should exist with measures of related concepts or scores produced in similar or diverse patient groups
  - Convergent/discriminant: Strength of correlation testing
  - Known groups: Degree the instrument can distinguish among groups

Measurement Properties: Reliability

- Test-retest (and intra-interviewer): Stability of scores over time when no change is expected in the concept of interest
- Internal consistency: Extent that items contributing to a score measure the same concept
- Inter-interviewer: Agreement among responses when administered by two or more different interviewers

Measurement Properties: Ability to detect change

- Evidence that a PRO instrument can identify differences in scores over time in patients similar to those in the clinical trials who have changed with respect to the measurement concept

Measurement Properties: Documentation for construct validity, reliability, and ability to detect change

- Protocols
  - Sample size, inclusion, exclusion criteria
  - Research design
- Statistical analysis plan
  - *A priori* hypotheses
- Study summaries
  - Testing results for each domain or summary score proposed as an endpoint to support claims
  - Proposed responder definition based on anchor-based approach

Translation and Cultural Adaptation

- Best addressed early in instrument development with either
  - Concurrent development in targeted language/culture groups
  - Participation by a linguistic expert to facilitate ease of translation of final instrument
- Cultural equivalence includes consideration of whether disease, standard of treatment, or measurement concept are the same or differ across language/culture groups

Translation and Cultural Adaptation: Documentation

- Methodologies used to achieve equivalence between source and target version of instrument
- Summaries of the following steps, if applicable:
  - Definition of concepts of original instrument
  - Translation from source language into target language
  - Translation from target language into source language
  - Acceptability test of formats used in target language
  - Comparison across source and target versions of cultural adaptation issues

PRO Instrument Manual

- Exact version of the instrument
  - All versions included
  - Screen shots of electronic format
  - Instructions for patient use; patient training
- Investigator training methods and materials
- Instrument administration guidelines
- Scoring algorithm

Qualification: Initial FDA review experience

- It's all new!
  - Substantial internal training
  - Support for qualification concept so far
- Variability in submissions
  - Guidance forthcoming
  - Submit using electronic media—NOT email
- Timing is everything
  - IBS Scoping Document review complete
  - IBS draft guidance publishes TODAY!
- EMA collaboration and harmonization encouraging

Scoping Document: Review Issues

- Composite instruments
  - Hypothesized conceptual framework of the instrument needs to represent whether….
    - a total composite score will be sought versus developing domain scores only
    - each item of a composite measure will be developed to support specific labeling claims versus domain/total scores only
- Keep in mind that hypothesized conceptual frameworks can change as more information is obtained during instrument development