Development of Validated Instruments

Michelle Tarver, MD, PhD
Medical Officer
Division of Ophthalmic and Ear, Nose, and Throat Devices
Center for Devices and Radiological Health
Food and Drug Administration

No Financial Conflicts to Disclose

Overview

- Define Patient Reported Outcomes (PROs)
- Factors to Consider when Developing PROs
- FDA Guidance for PROs
- Use of PROs in FDA Clinical Trials

Patient Reported Outcomes (PROs)

- 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
- Can be measured in absolute terms (e.g., severity of a symptom) or as a change from a previous measure
- In trials, measures the effect of a medical intervention on one or more concepts
  - Concept is the thing being measured (e.g., symptom, effects on function, severity of health condition)

Concepts a PRO May Capture

- Symptoms
- Symptom impact and functioning
- Disability/handicap
- Adverse events
- Treatment tolerability
- Treatment satisfaction
- Health-related quality of life

Criteria to Consider in PRO Development

- Appropriateness
  - Does the content address the relevant questions for the device?
- Acceptability
  - Is the questionnaire acceptable to patients?
- Feasibility
  - Is it easy to administer and process/analyze?
- Interpretability
  - Are the scores interpretable?

Abstracted from (1) Patient Reported Outcomes Measurement Group: University of Oxford
(2) NIH PROMIS Instrument Development and Validation Standards

Criteria to Consider in PRO Development

- Precision
  - How precise are the scores?
- Reliability
  - Does it produce results that are reproducible and internally consistent?
- Validity
  - Does the questionnaire measure what it claims to measure?
- Responsiveness
  - Does the questionnaire detect changes over time that matter to patients?

Criteria to Consider: Appropriateness (Concepts Measured)

- Conceptual framework to support measurement of the concept of interest
  - Patient interviews, focus groups, qualitative cognitive interviewing inform this process
  - Will evolve with acquisition of empirical data
- Recall period appropriate for population, disease state, or application of the questionnaire

Criteria to Consider: Acceptability

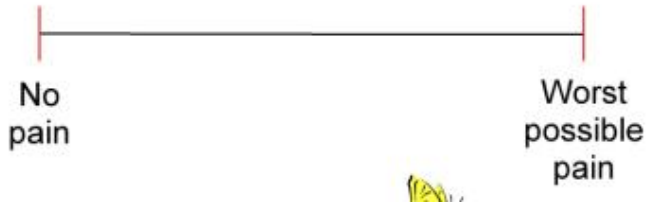
- Administration
  - Mode (self vs. interviewer; paper vs. computer)
  - Time (length of time it takes to complete; frequency of administration)
  - Format of the instrument (layout, appearance, legibility)
  - Language (considering associated idioms and cultural norms)
  - Costs (training of staff, printing of questionnaires, electronic devices)

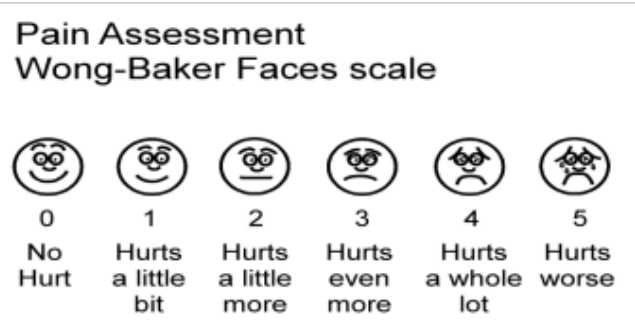
Criteria to Consider: Interpretability

- The meaningfulness of scores produced by the questionnaire
  - What does a score mean?
  - What is the minimal clinically important difference (minimum score changed deemed beneficial to patients)?
  - Should an overall score be computed and presented to support a given claim?

Criteria to Consider: Precision

- How is the questionnaire scaled?
  - Binary (yes vs. no)
  - Likert/adjectival (e.g., strongly agree, agree, disagree, strongly disagree)
  - Visual analogue
  - Pictoral
  - Weighting of items
- Are there large floor or ceiling effects on the score?
  - Limits discriminatory power and responsiveness





Criteria to Consider: Reliability & Reproducibility

- Internally consistent or reproducible and degree to which questionnaire is free from measurement error
  - Proportion of score that is signal rather than noise
  - As measurement error increases ≥ sample size increases to obtain precise estimates of intervention effect
  - Target 0.7-0.9 in many studies
- Questionnaire yields same results on repeated administrations without any intervention
  - Test-retest reliability (e.g., 2-14 days)

Criteria to Consider: Validity

- Extent to which a questionnaire measures what is intended
  - Qualitative research with the targeted patient population is needed to ensure developmental appropriateness of the measure
    - » Face validity (appears to measure concept of interest)
    - » Content validity in context of use (adequately covers concept/domain of interest)
  - Quantitative
    - » Criterion validity (correlates with another measure considered more accurate. May or may not be available)
    - » Construct validity

Establishing Content Validity

- Literature review
- Expert opinion
- Qualitative research (essential)
  - Input from target population of patients to document understandability and comprehensiveness of measure
  - Diversity in demographic & disease characteristics of target population
- Quantitative analyses
  - Rasch
  - Factor analysis
  - Does not eliminate need for high quality cognitive debriefing of the final instrument in the relevant patient population

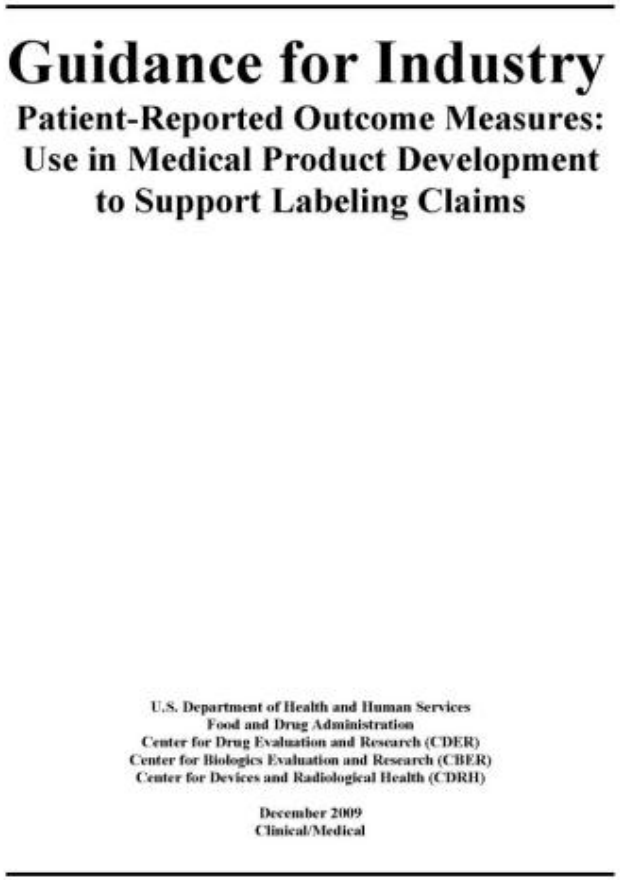
Establishing Content Validity (cont.)

- Determined after confirmation that the concept and the context of use are appropriate
- Empirical evidence that the instrument measures the targeted concept in the context of use
  - If existing instrument is used for a new context of use, additional content validity evidence may need to be developed
- Content validity must be established before other evidence of construct validity, reliability or sensitivity to change can be interpreted

Criteria to Consider: Responsiveness

- Captures health changes
  - before and after the intervention OR
  - in different disease or treatment states
- Evaluated within specific populations and not a fixed/inherent property of the questionnaire
- Determine the relevant, clinically meaningful effect size

FDA Guidance on PROs



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

- FDA-wide guidance
- Acknowledges the importance of appropriately and effectively incorporating the patient's voice into the evaluation of medical products
- Final: December 2009

http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM205269.pdf

PRO Development is Iterative

- Reasons for changing items during development
  - Clarity or relevance
  - Response range
  - Variability
  - Reproducibility
  - Inter-item correlation
  - Ability to detect change
  - Item discrimination
  - Redundancy
  - Recall period

http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM205269.pdf

Outcome Measure

- The impact of treatment on how patients see, feel, and function in their daily lives
  - Must be well-defined and reliably measured
  - Can be assessed directly (e.g., visual symptoms)
  - Can be assessed indirectly (e.g., visual acuity)

FDA Review of Clinical Trial Outcome Assessments

- Identify the measurement concepts
  - Does the instrument measure the *concept* it was intended to measure?
  - Does the instrument measure the *concept* claimed?
- Identify the context of use
  - Primary or secondary endpoints?
  - Trial inclusion criteria?

http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM205269.pdf

Designing Clinical Trial Consider use of PROs

- Step 1: Define the diseased population
- Step 2: Define the context of use
- Step 3: Select concepts of measurement that will define treatment benefit or safety concern

Designing Clinical Trial Consider use of PROs (cont.)

- Step 4: Select or develop a well-defined and reliable outcome assessment to measure each concept for the proposed context of use
  - If not observable, need a PRO
  - Observable but does not need clinical judgment may need PRO as well
  - Self-report of symptoms provides direct evidence of treatment benefit or harm and should be used whenever possible

Elements of PRO Submissions Reviewed by FDA

- Concepts being measured
- Number of items
- Conceptual framework of the instrument
- Medical condition and population for intended use
- Data collection method
- Administration mode
- Response options
- Recall period
- Scoring (weighting of items or domains)
- Format
- Translation or cultural adaptation availability

http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM205269.pdf

Measurement Properties of PRO

- Should be well established before enrollment in pivotal clinical trial
- Requests for FDA input
  - Need information about labeling goals
  - Hypothesized PRO instrument conceptual framework
  - Relationship of PRO endpoints to clinical trial

Summary

- PROs should be
  - Valid with respect to content in the intended population
  - Psychometrically evaluated
  - Well-developed before utilization in the pivotal clinical trial
  - Considered during the design and planning of the clinical trial
  - Developed with early FDA input

Thank you