



Clinical trial endpoints: Development and validation of measures to support claims in labeling

Accelerating Therapies for Rare Diseases Workshop

Laurie Burke

Associate Director for Study Endpoints and Labeling
Office of New Drugs, CDER, FDA

October 19, 2010

Presentation Outline

- Review of rating scales
 - Regulations
 - Key considerations
- Drug development tool qualification
 - What is “qualification?”
 - FDA process

Treatment Benefit

- The impact of treatment on how a patient ***survives, feels, or functions***
- Always comparative
- Can be related to either effectiveness or safety
 - Improvement or delay in the development of disease-related symptoms
 - Reduction or delay in treatment-related toxicity

Claim

- Any statement of treatment benefit
- Can be found anywhere in FDA-approved labeling or in advertising
- By regulation:
 - Supported by substantial evidence
 - **Not** false or misleading

Substantial Evidence

- 21 CFR 314.126(a)
 - *“**Reports of adequate and well-controlled investigations provide the primary basis for determining whether there is ‘substantial evidence’** to support the claims of effectiveness for new drugs.*
 - *Therefore, the study report should provide **sufficient details of study design, conduct and analysis** to allow critical evaluation and a determination of whether the characteristics of an adequate and well-controlled study are present.”*

Characteristics of an *Adequate and Well-Controlled Study*

- 21 CFR 314.126 (b)
 - (1) Clear statement of objectives
 - (2) Study design permits valid comparison (appropriate control)
 - (3) Select patients with disease/condition (treatment) or at risk of disease (prevention)
 - (4) Baseline comparability (randomization)
 - (5) Minimize bias (blinding, etc.)
 - (6) **Appropriate methods for assessment of outcome**
 - (7) Appropriate methods of analysis

Appropriate Methods for Assessment of Outcome

- 21 CFR 314.126(b)(6)

*“The methods of assessment of subjects’ response are **well-defined and reliable**. The protocol for the study and the report of results should **explain the variables measured, the methods of observation, and the criteria used to assess response.**”*

Endpoint

- The measurement that will be statistically compared among treatment groups to assess the effect of treatment and corresponds with the clinical trial's objectives, design, and data analysis plan
- Endpoint measure requirements
 - Well-defined and reliable
 - Meaningful
 - Interpretable
 - “Fit for Purpose”

Types of Endpoint Measures

- **Objective Test** (e.g. weight, hemoglobin)
- **Patient Reported Outcome (PRO)**
 - any report of the status of a patient's health that comes directly from the patient (e.g. pain intensity, chronic heart failure symptoms)
- **Observer Reported Outcome (ObsRO)**
 - observations but not interpretation of a patient's health status (e.g. parent assessment of baby's vomiting episodes)
- **Clinician Reported Outcome (ClinRO)**
 - Clinician rating of a patient's status or condition (e.g. vertebral fracture, clinical cure of pneumonia)

Goal of Measurement: “Concept”

- **Concept:** The specific goal of measurement (i.e., the thing or event that is to be measured)
 - Forms the basis for describing claims in labeling
- Can represent:
 - Disease-specific symptoms and signs
 - Impact of signs/symptoms (e.g., on activities of daily life)
 - Impact of disease on health related quality of life

Instruments and Rating Scales

- **Instrument:**

- A means to capture data plus all the information and documentation that supports its use

- **Rating Scale:**

- A type of instrument that assigns a rating to a concept to produce a score
- Can be patient-reported, observer-reported, or clinician-reported

Example: 11-point numeric rating scale of pain intensity

FDA PRO Guidance

Published: December 2009

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

- Explains how FDA reviews whether a PRO instrument is “well-defined and reliable” and adequate to support medical product claims
- FDA’s conclusions are based on the proposed context of instrument use as a clinical trial endpoint measure
- PRO guidance principles apply to review of other types of rating scales:
 - Clinician rating scales
 - Observer rating scales

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

Instrument Review Considerations

- Medical Condition for Intended Use
- Concept(s) being measured
- Framework (e.g., subscales)
- Instructions for use
- Items
- Response options
- Recall period
- Scoring
- Guidelines for interpretation of change
- Cultural adaptation/translation
- Measurement property documentation

Measurement Properties

- **Content Validity**
- **Construct Validity**
- **Reliability**
 - Test-retest
 - Internal Consistency
- **Ability to detect change**

Content Validity

- Evidence that the instrument measures the concept of interest including evidence that the items and domains of an instrument are **meaningful, comprehensive, appropriate, and interpretable**, relative to its intended context of use
- **Critical for interpretation and labeling**
- **Should be established prior to evaluating other measurement properties**

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 clinical trial study population
 - Measurement concept conforms with the proposed clinical trial objectives

Content Validity Evidence: Qualitative and Quantitative Research

- Literature review and expert input on concept, disease, existing instruments
- Protocol-driven
- Qualitative research in the respondent population
 - Concept elicitation interviews
 - Cognitive testing before finalization of instrument content
 - Qualitative research design should be adapted for small sample (rare disease)
 - Evidence of concept saturation
 - Concept confirmation/testing of the questionnaire might be combined
- Statistical exploration of qualitative data
 - E.g., Text analysis software, factor analysis, item response analysis

Instrument Implementation Review

- Study design (blinding and randomization)
- Enrollment criteria
- Endpoint model
 - Hierarchy of endpoints that support claims (e.g. primary, secondary)
- Frequency of outcome assessments
- Clinical trial duration
- Responder analysis (if applicable)
- Statistical analysis plan
 - Methods for handling missing data (*a priori*)

Outcome Considerations in Rare Disease Clinical Trials

- Small sample size
 - Sensitive outcome measures are advantageous
- Availability of adequate outcome measures
 - Well-defined and reliable
 - Responsive to change
 - Interpretation of a clinically meaningful effect
- Validity of Outcome Data
 - Unintentional unblinding (clinician and/or patient) due to toxicity or drug product

Outcome Considerations in Rare Disease Clinical Trials (cont)

- Pediatric measures
 - Appropriate age for reliable self-report
 - Relevant concepts for pediatric patients of given age range
 - Concurrent caregiver assessments

- Multinational trials
 - Translation and cultural adaptation

Pending Draft Qualification Guidance

- Covers qualification for drug development tools (DDTs):
 - Biomarkers
 - PRO tools and other rating scales
 - Others
- Defines qualification
- Describes the qualification process
- Refers to existing evidence standards (e.g., regs and guidance's)
- New DDTs *and* existing DDTs

DDT Qualification: What it is

- Regulatory conclusion that within the stated context of use, the results of drug development tool (DDT) measurement can be relied upon to have a stated interpretation and utility— “fit for purpose”
- Qualification is independent of a specific application
- For rating scales:
 - Data produced can be interpreted as clinically meaningful and can be used as a primary or key secondary endpoint to support a claim in labeling

DDT Qualification: Rationale

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

DDT Qualification: Process

- FDA reviews Scoping Stage Summary Document before agreeing to participate in Qualification Stage 1
- DDT developer provides background information so FDA can decide whether or not to participate in the qualification process for the DDT
- A Scoping Stage Summary Document includes:
 - Targeted endpoint model
 - Targeted claim
 - Hypothesized conceptual framework
 - Rationale for DDT qualification
 - Methodology to collect evidence to support DDT qualification

Rating Scale Qualification: Process

- *Qualification Stage 1: Consultation and advice*
 - FDA convenes a Qualification Review Team to review and comment at certain times during development

- **Content validity summary**

{ e.g., qualitative research protocols (concept elicitation, cognitive debriefing), transcripts, reports, draft of the instruments, revised CF, translation methodology, item map

- **Scoring development summary**

{ e.g., protocols for item reduction, analyses plan, scoring algorithm, draft of the instruments, item map

- **Instrument finalization summary**

{ e.g., protocols, analyses plan, report, final instrument, item map, interpretation

DDT Qualification: Process

- *Qualification Stage 2: Review and qualification decision*
 - Once DDT is ready for qualification, the Agency will review a **Qualification Package** (“dossier”)
 - Public notice of DDT qualification, as appropriate
 - Public availability of DDT tool

Conclusions

- All endpoint measures need to be well-defined and reliable to support claims in labeling
 - Patient-reported
 - Observer-reported
 - Other
- Many diseases are still without adequate clinical trial endpoint measures to demonstrate efficacy
- FDA qualification may improve the efficiency of endpoint measure development
- Well-defined and reliable endpoint measures are particularly important in rare diseases where sample sizes are limited