Treatment Benefit and Clinical Outcome Assessment

Laurie Burke, Director
Study Endpoints and Labeling Development Staff
Office of New Drugs, CDER, FDA

FNIH-NIA-FDA Sarcopenia Consensus Summit,
Foundation for the National Institutes of Health
May 11, 2012 Baltimore
Overview

• Definition of treatment benefit
• Identification or development of adequate clinical outcome assessments (COAs)
  – Nomenclature
  – Context of use
  – Concept of measurement
  – Content validity
• Advantages of “well-defined and reliable” assessments in clinical trials
Treatment Benefit

• The impact of treatment on how a patient “survives, feels, or functions” daily life
  – Measured as effectiveness or comparative safety
  – Used interchangeably with “clinical benefit”
  – Can be measured directly or indirectly

• Described in labeling as a claim using words that represent the concept measured by the COA in the context of use defined by the study protocol
Nomenclature for Universe of Efficacy Endpoint Measures

• Clinical Outcome Assessments (COAs)
  – Patient reported outcome assessments (PROs)
  – Clinician reported outcome assessments (ClinROs)
  – Observer reported outcome assessments (ObsROs)

• Biomarkers
  – Assessment is not influenced by human choices
  – Most often intended to predict a future state of survives, feels or functions

• Survival
COAs
PRO, ClinRO and ObsRO Assessments

• All influenced by human choices
  – Conscious or unconscious

• Clinician, observer, or patient rater
  – Judgment, cooperation, motivation involved

• Must be used to directly assess “feels or functions”
  – Only patients can assess symptoms

• Many can be used to indirectly assess “feels or functions” (as a replacement for direct assessment)
  – Not all patients can rate themselves
  – Not all functioning can feasibly be measured directly
Selecting/Developing a COA

Step 1: Define context of use

Step 2: Define concept(s) of measurement and reporter

Observable Concept
- No Clinical Judgment Needed
- Self-report? No
- ObsRO
- PRO
- ClinRO

Non-Observable Concept
- Clinical Judgment
- PRO

Physiologic or lab findings that can be measured without human assessment
- Biomarker

Step 3: Document evidence that the content of the COA is a valid assessment of the concept of measurement in the context of use

Step 4: Document evidence that the other measurement properties of the instrument are adequate in the context of use
Context of Use: Define the Disease

• Disease definition should be explicit and specific to targeted clinical trial population
  – Inclusion/exclusion criteria for study protocol
• Disease definitions for clinical practice are generally too general for clinical trials
• The disease definition may be variable according to age group
  – May need group-specific disease definitions
Other Aspects of Context of Use

• Specific target population-the enrollment criteria for the study population including:
  – Screening criteria (e.g., disease characteristics)
  – Demographics (don’t forget the kids!)
  – Cultural and language considerations

• Clinical setting (e.g., inpatient vs. outpatient)

• General plan for study design (study objectives, endpoint model)

• General plan for data interpretation

• Targeted labeling claims
## Context of Use: Endpoint Model

An Endpoint Model displays the role and hierarchy of relevant outcome concepts in clinical trials (i.e., all primary and secondary endpoints).

<table>
<thead>
<tr>
<th>Endpoints</th>
<th>Concepts</th>
<th>COA/Biomarker/Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>Concept A</td>
<td>OA 1</td>
</tr>
<tr>
<td>Secondary with</td>
<td>Concept B</td>
<td>OA 2</td>
</tr>
<tr>
<td>Hierarchy</td>
<td>Concept C</td>
<td>OA 3</td>
</tr>
<tr>
<td></td>
<td>Concept D</td>
<td>OA 4</td>
</tr>
<tr>
<td></td>
<td>Concept E</td>
<td>OA 5</td>
</tr>
<tr>
<td>Exploratory</td>
<td>Other concept</td>
<td>Other OA</td>
</tr>
</tbody>
</table>
Concept of Measurement
Direct Evidence of Tx Benefit (Proximal to Distal)

Disease-defining concepts

Proximal disease Impact concepts

Distal disease Impact concepts

Disease impact on general life concepts

Core signs, symptoms or decrements in functioning

Related functioning

Additional functioning

General psychological functioning

Productivity

Health status

Health-related quality of life

General physical functioning

Social functioning

Satisfaction with health

Related S/Ss

Additional S/Ss
Indirect Assessment of Treatment Benefit

• Indirect means the concept being measured is not the directly meaningful concept—it’s a replacement

• “Indirectness” is relative
Outcome Assessments Vary in Their Relationship to Direct Evidence of Treatment Benefit

Survive, Feel, Function in normal daily life → Direct

Functional assessments in clinic → Strength testing

Bio-markers → Indirect

Indirect COA assessment needs empiric justification for replacement value and relationship to how patients survive, feel or function
FDA Recommendation for a COA

• Specific to a context of use
• Specific to a version of an instrument including mode of administration and training materials (ETDRS Eye Chart)
• Specific to the concept of measurement (visual acuity)
• For indirect assessments of treatment benefit, specific to the meaningful aspect of “feels or functions” indirectly measured (vision-dependent activities)
Documenting COA Measurement Properties

- Defines how the Agency interprets “well-defined and reliable” (21 CFR 314.126) for PRO measures intended to provide evidence of treatment benefit
- Summarizes good measurement principles applicable to any PRO, ClinRO or ObsRO assessment

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

Well-defined and Reliable:
MEASUREMENT PROPERTIES found adequate to measure the CONCEPT OF MEASUREMENT in the CONTEXT OF USE

• Measurement Properties:
  – Content validity
  – Construct validity
  – Reliability (particularly test-retest)
  – Ability to detect change
    • Information to support interpretation of change
Content Validity

- Empiric evidence that the instrument measures the targeted concept in the context of use
  - If existing instrument is to be adapted for a new context of use, additional content validity evidence may need to be developed

- Established before evidence of construct validity, reliability or sensitivity to change can be interpreted
  - Construct validity provides only circumstantial evidence
  - There’s no such thing as a “validated instrument.” FDA reviews content validity in each context of use.
Methods to Establish Content Validity Are Iterative

- Literature review
- Expert opinion
- Qualitative research
  - Input from target responder population to document understandability and comprehensiveness
    - PRO: target population of patients
    - ObsRO: target population of respondents
    - ClinRO: target population of clinicians
- Quantitative exploration (e.g., Rasch Measurement Theory)
  - Item characteristics (their cohesiveness as a group, their ability to map a variable in a clinically meaningful way, their stability and bias)
  - Item response options (do they "work" empirically as intended conceptually)
  - Initial understanding of score interpretation
Quantitative Support for Content Validity

- **Unidimensionality**: score represents a single concept
- **Hierarchical order**: scale represents less severe to more severe
- **Equal spacing**: response options are correctly ordered and spaced from less severe to more severe
Other Measurement Properties

• After content validity is established, longitudinal studies establish other measurement properties:
  – Construct validity
  – Reliability
  – Ability to detect change

• Need to be demonstrated with the final version of the instrument to estimate an effect size and interpret clinical trial results
Measures to Provide Indirect Evidence of Treatment Benefit

- Still need to demonstrate that COA measures the indirect construct in a “well-defined and reliable” way
- In addition, need to demonstrate the relationship of the indirect construct to how patients survive, feel or function to support the replacement value (often described as evidence of clinical meaningfulness)
Validity Impacts Variability, Assay Sensitivity and Interpretability

• Patient heterogeneity identified and addressed
  – Validity considered in all important population subsets in context of use

• Random variation (error) minimized
  – Test-retest or inter-rater reliability optimized in context of use

• Systematic, non-random variation minimized
  – Validity tested in all population subsets in context of use

• Experiment design and conduct error minimized
  – Explicit training and instructions specific for context of use
Conclusion

• An assessment cannot be chosen or developed without a well-defined context of use and targeted concept of measurement

• Rigor in measurement and standardization across clinical trial investigative sites is critical to conserve clinical trial resources

• The science of measurement continues to evolve with new tools and methods for efficient development and modification of assessments