



An FDA Perspective on Clinical Trial Endpoint Measurements

ACTION-II

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

26 October 2012 – White Oak



Treatment Benefit

- The impact of treatment on how a patient “survives, feels, or functions” in daily life
 - Measured as effectiveness or comparative safety
 - Used interchangeably with “clinical benefit”
 - Can be measured **directly** or **indirectly**
- Described in labeling or promotion as a claim that describes the benefit measured in the context of use defined by the study protocol

Direct vs. Indirect

Treatment Benefit Assessment

- Survival
- Symptom (e.g., pain)
- Function (e.g., physical function)

Direct Assessments



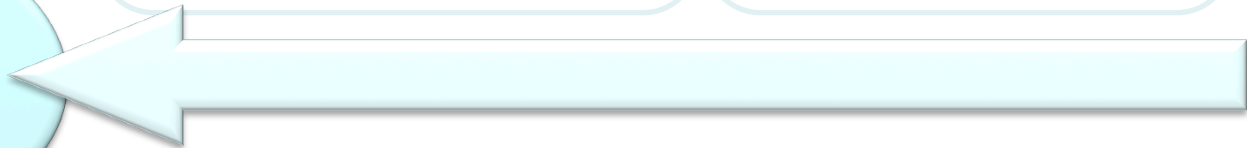
Treatment Benefit

(tx impact on how patients feel, function, survive)

- 6-Minute walk test
- Pulmonary function (e.g., FEV1)
- Use of rescue pain medication

- HbA1c
- Prostate specific antigen (PSA)
- Tumor size
- Other biomarkers

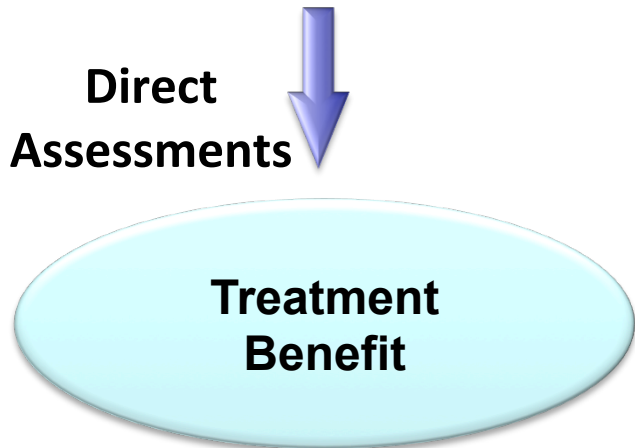
Indirect Assessments



Direct vs. Indirect Treatment Benefit

Example: Asthma

- Symptoms (e.g., cough, shortness of breath)
- Nighttime awakenings/ sleep disturbance
- Activity Limitation

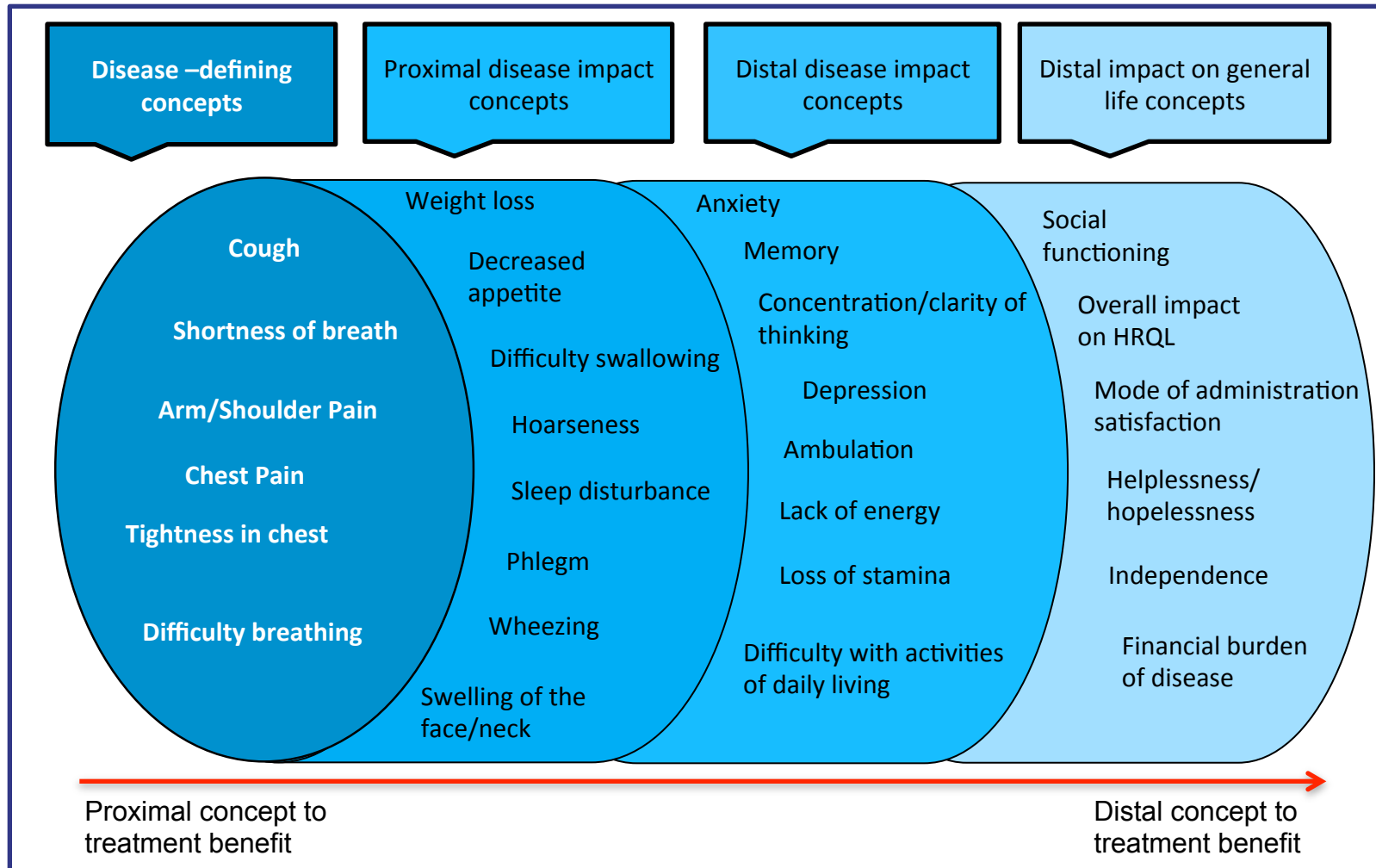


- Lung Function e.g.
 - FEV1(%predicted)
 - PEF
- Airway hyperresponsiveness

- Exhaled Nitric Oxide (Fe_{NO})
- Sputum Eosinophils
- Blood Eosinophils



Direct Evidence of Treatment Benefit: Non-Small Cell Lung Cancer*



* Courtesy of the PRO Consortium, Critical Path Institute. Concepts identified through a cursory review of the literature. Graph will evolve based on findings from qualitative research and clinician expertise



What is “Indirect” Evidence of Treatment Benefit?

- The concept being measured is different from the directly meaningful concept—it’s a replacement
- “Indirectness” is relative and permanent
- FDA reviews evidence of the relationship between the indirect concept and how patients survive, feel or function
 - Generally requires longitudinal studies to demonstrate that relationship
 - Often described as evidence of clinical meaningfulness



Types of Outcome Assessments in Clinical Trials

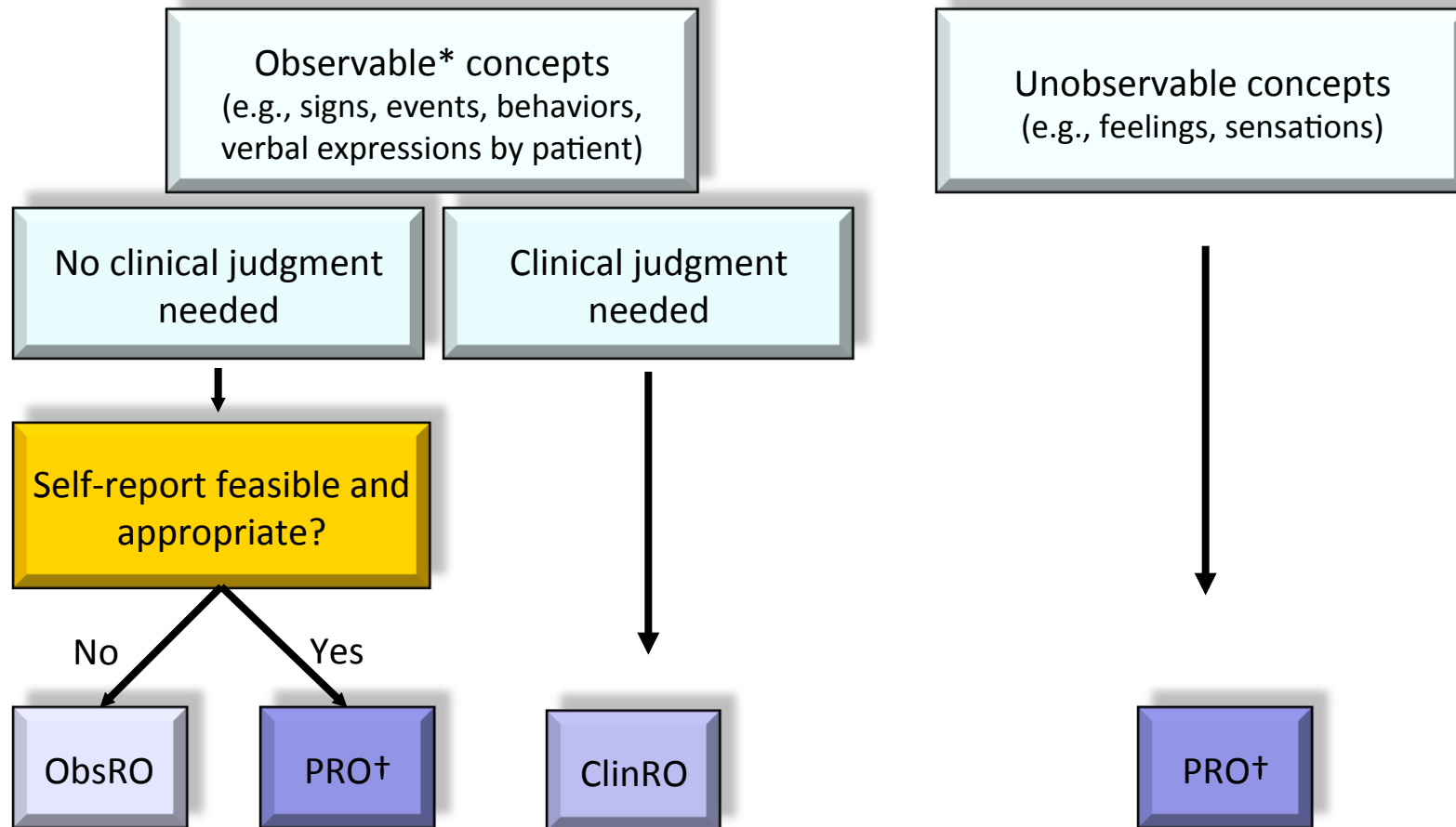
- **Clinical Outcome Assessments (COAs)**
 - “Reported” assessments (subject to influence by human choices, judgment, cooperation, or motivation)
 - Patient-reported outcome assessments (PROs)
 - Clinician-reported outcome assessments (ClinROs)
 - Observer-reported outcome assessments (ObsROs)
- **Biomarkers**
 - Results not influenced by humans; relies on a standardized, automated process
- **Survival**



PRO, ClinRO and ObsRO Assessments

- All influenced by human choices
 - Conscious or unconscious
 - Judgment, cooperation, motivation involved
- Most ClinROs and ObsROs and all biomarkers **indirectly** assess “feels or functions” (as a replacement for direct assessment)
- Indirect assessment often necessary because
 - Not all patients can rate themselves
 - Not all functioning can feasibly be measured directly in normal daily life

Observable vs. Unobservable Concepts



* Observable concepts: must be able to be detected by a sense or senses -- vision, hearing, smell, or touch

† PRO instrument could be self-completed or interview-administered



Review of PRO, ClinRO, and ObsRO Measurements

Guidance for Industry

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

[http://www.fda.gov/
downloads/Drugs/
GuidanceComplianceRegulat
oryInformation/Guidances/
UCM205269.pdf](http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM205269.pdf)

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

- 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



DDT Qualification Process Guidance (Draft)

Guidance for Industry

Qualification Process for Drug Development Tools

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact (CDER) Shaniece Gathers, 301-796-2600.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

October 2010
Clinical Medical

- Qualification process for drug development tools (DDTs):
 - Biomarkers
 - Clinical outcome assessments (PRO, ClinRO and ObsRO measures)
 - Animal models
 - May be others in future
- New *and* existing DDTs
- Not required for tool use in drug development
- Guidance emphasis is on *process*

CDER DDT website:

[http://www.fda.gov/downloads/Drugs/
GuidanceComplianceRegulatoryInformation/Guidances/
UCM230597.pdf](http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM230597.pdf)



What Is Qualification?

- A conclusion that within a carefully and specifically stated “context of use,” a drug development tool has been demonstrated to reliably support a specified manner of interpretation and application in drug development
 - Utility in drug development, particularly clinical trials, is central to purpose of qualification
 - Particularly intended for tools expected to have application in multiple different drug development programs



What Biomarkers Are Eligible for Qualification?

- Prognostic biomarker
 - Indicates future clinical course of the patient with respect to some specified clinical outcome, in the absence of a Tx intervention
- Predictive biomarker
 - Measured prior to an intervention
 - Identifies patients who are relatively susceptible to a particular drug effect versus less susceptible patients
- Pharmacodynamic biomarker
 - Response-indicator biomarker
- Efficacy-response biomarker
 - Efficacy-surrogate biomarker, Surrogate endpoint
 - Subset of general pharmacodynamic biomarkers



What COAs Are Eligible for Qualification?

- Measurement is intended to support primary or secondary endpoints related to treatment benefit
- Context of use is well-defined
- Measurement concept is appropriate for the context of use
- Review of evidence provides confidence that the assessment adequately measures the measurement concept and the evidence is specific to:
 - **The concept of measurement**
 - **The context of use**



Context of Use Includes....

- Disease definition
 - Explicit and specific to targeted clinical trial population
 - Matches the inclusion/exclusion criteria
 - Disease severity
 - Demographics
 - Other important aspects of heterogeneity
 - More detailed than diagnostic or stratification criteria
 - May vary by subgroup (e.g., age)
- Clinical setting (e.g., inpatient vs. outpatient)
- General plan for study design (study objectives, endpoint model)
- General plan for data interpretation
- Targeted labeling claims (consistent with the concept of measurement)



Natural History Studies

- Can provide the basis for describing the disease
 - Track course of disease over time
 - Provide information about variability/heterogeneity
 - Identify demographic, genetic, environmental and other variables that correlate with disease and outcomes in the absence of treatment
- Contribute to scientific foundation upon which drug development programs can be built
- Independent of individual investigational agents
- Most informative when NH study data are available early in development Ideally before design of efficacy trials
- Patient and caregiver involvement is important
 - Engage all stakeholders early and on an ongoing basis
- Institute of Medicine. 2010. *Rare Disease and Orphan Products. Accelerating Research and Development*



Develop Evidence of Validity First

- Empiric evidence that the score is a measurement of the intended concept in the specified context of use
 - Claims in labeling are based on confidence that the claimed concept (direct or indirect evidence of treatment benefit) was measured validly and results were interpretable in the context of use studied
 - Traditional statistical tests of validity (internal consistency, correlations with other measures, known group differences) do not tell us what a score represents
- Established before evidence of construct validity, reliability or sensitivity to change can be interpreted



What Is the Relationship between Content Validity and Context of Use?

- Content validity is specific to the context of use in which the evidence was generated
- If the existing measurement is to be adapted for a new context of use, additional content validity evidence may need to be developed
- FDA reviews content validity within each context of use
 - There's no such thing as a “validated instrument.”



Initial Research to Develop Valid COAs

Qualitative

Includes literature review and expert opinion

- Protocol and hypothesis driven
- Input from target responder population to document understandability and comprehensiveness (interviews, focus groups)
- Multiple rounds of qualitative research necessary to support
 - Development of the content
 - Refinement of the content
 - Confirmation of validity with the final content and in the final format

Quantitative

Can be used iteratively with qualitative evidence to finalize measurement content

Includes evidence that:

- Scores represent a single concept
- Scale represents less severe to more severe
- Response options are correctly ordered and spaced from less severe to more severe
- The range and distribution of scores is adequate for the context of use?

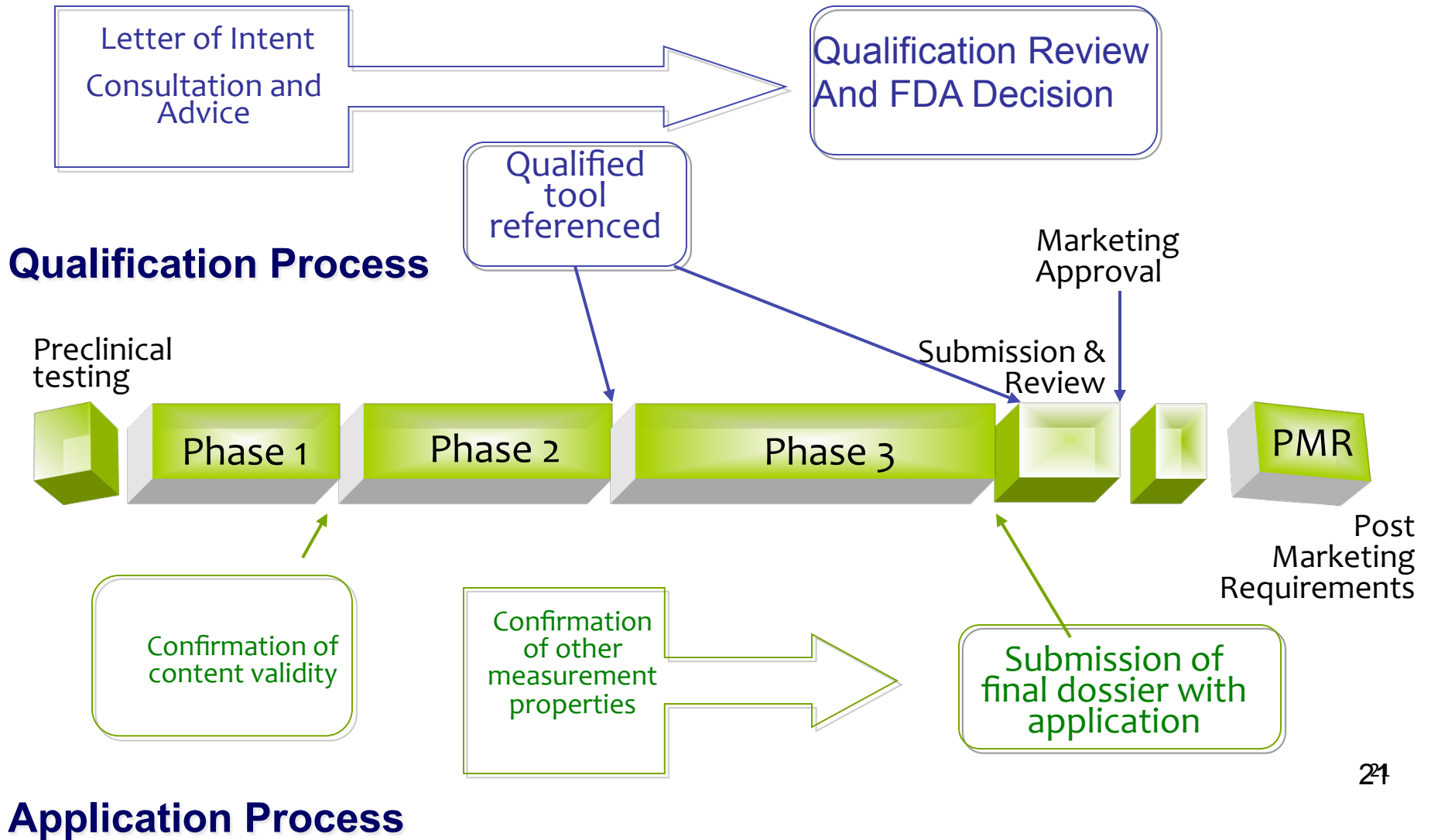


Completing the Dossier

- Longitudinal studies to establish other measurement properties:
 - Construct validity (if a hypothesized relationship to other measures exists)
 - Reliability (e.g., test-retest in stable patients)
 - Ability to detect change
 - Mean change from baseline
 - Responder definition
- All need to be demonstrated with the final version of the instrument



FDA Review of COAs for Use in Drug Development: Two Processes





Review of a COA Measurement

- What concept is represented by the score?
- What is the context of use proposed?
- Is the concept clinically meaningful/important/relevant in the proposed context of use?
- Does the concept directly measure how patients feel and function in daily life?
- If not, what is the replacement value of the measure?
- Does the COA measure the concept in the appropriate context of use in a valid way?
- What are its other measurement properties (traditional validation data) in the context of use?