Clinical Trial Endpoints

Eugene J. Sullivan, MD FCCP
Principal
EJS Consulting, LLC

Orientation

Establishing Efficacy

- ✓ Legal / Regulatory Basis
- ✓ What and How
  - "substantial evidence"
  - "adequate and well-controlled" trials
    - trial designs, types of controls
- "Substantial Evidence" of What?
  - Appropriately designed and analyzed trials can establish an effect of the drug on something. What should that something be?
- [This discussion will focus on primary endpoints for Phase 3 trials. Development program should also result in data to provide *adequate directions for use*]
  - dose response, demographic subgroups, pharmacokinetics, drug-drug interaction, etc.

The Law Says…

- A marketing application will be rejected if there is “a lack of substantial evidence that the drug will have the effect it purports or is represented to have … in the proposed labeling..”
- So why doesn‘t the FDA approve any drug, as long as the labeling truthfully states what effect has been demonstrated?

Answer: The effect must be clinically meaningful.

(Not in the Act, but established by Warner-Lambert v Heckler, 1986)

“Feels, Functions, or Survives”

- All drugs have safety risks. Therefore, the only reason that a patient would want to take a drug would be if the drug:
  - resulted in a benefit that was detectable by the patient (improvement in symptoms, improvement in functional capacity),
  - improved survival

  or
  - decreased the chances of developing a condition or disease complication that is itself apparent to the patient and is undesirable (e.g. stroke)
- Therefore, a primary endpoint should be a direct measure of one of these. A primary endpoint should generally not be a measure of something that is not important to the patient (exception: validated surrogate endpoint).

“Direct” Endpoints

- Clinically meaningful endpoints that directly measure how a patient **feels**, **functions**, or **survives**
- Endpoints that in themselves represent or characterize the clinical outcome of interest
  - Objective: survival, disease exacerbation, clinical event (e.g. MI, stroke), etc.
  - Subjective: symptom score, “health related quality of life” (validated instrument), etc.
- Customarily, the basis for approval of new drugs

Note: The term “direct” is used here to distinguish from “surrogate” endpoints, but this term is not uniformly utilized. Others may refer to these as “true” or “clinically meaningful” endpoints

Surrogate Endpoints

- A surrogate endpoint is a laboratory measure / physical sign that is intended to be used as a substitute for a clinically meaningful endpoint.
- Changes induced by a therapy on a surrogate endpoint are expected to reflect changes in a clinically meaningful endpoint.
- This expectation must be supported by strong data (“validation”).
  - examples of failures of apparently reasonable proposed surrogate endpoints have led to caution (e.g. CAST Trial).
- Ideally, the surrogate should exist within the therapeutic pathway between the drug and meaningful benefit
  - i.e. the drug results in the therapeutic benefit by virtue of its effect on the surrogate

Surrogate Endpoints

- Surrogate endpoints can be used for drug approval:
  - if well validated, or
  - under Subpart H (21 CFR 314.500- 560; “accelerated approval” for serious and life-threatening illnesses; 1992)
    - requires adequate and well controlled trials
    - requires demonstrated effect on surrogate endpoint that is “reasonably likely, based on epidemiologic, therapeutic, pathophysiologic, or other evidence, to predict clinical benefit”
    - requires that the Applicant carry out, with due diligence, further adequate and well controlled studies to verify and describe the clinical benefit of the surrogate (where there is uncertainty as to the relation of the surrogate to the clinical benefit)

Why use a surrogate endpoint?

- Faster and easier to study
  - e.g., BP or cholesterol vs. survival, stroke, MI
- Faster drug development
  - get good drugs to patients sooner
- Cheaper
- Proving effect on direct endpoint may not be feasible
  - very low event rates
  - rare diseases

Examples of Surrogate Endpoints

- Hypertension
  - arterial blood pressure: surrogate for CVA, MI, heart failure
- Hypercholesterolemia
  - cholesterol levels: surrogate for atherosclerotic disease
- HIV
  - CD4 count or viral load: surrogate for complications of HIV
- Glaucoma
  - intraocular pressure: surrogate for loss of vision
- Diabetes Mellitus
  - blood glucose / hemoglobin A1c: surrogate for complications

Validating a Surrogate Endpoint

- For a surrogate to be useful, the relationship between the surrogate and the “direct” endpoint must be firmly established. Simple correlations, no matter how strong, are not enough.
- Ideal method: Analyses of multiple studies of known effective drugs, which assess both the direct and surrogate endpoints, in order to establish (and quantitate) the relationship.
- Very difficult to specify what type and quantity of data are sufficient to adequately validate a surrogate for use as a primary endpoint in a Phase 3 trial.
- Once validated, a surrogate may be useful for future studies of medicines, particularly those with same mechanism of action

Surrogate Endpoints: Potential Pitfalls

- Unless validated, the relationship between surrogate and direct benefit may not be causal, despite a strong association
  - e.g. fever / WBC as surrogate for treatment of pneumonia

- Drugs may have other unfavorable effects, apart from effect on surrogate.
  - e.g. if an antihypertensive lowered BP but also raised cholesterol, arterial blood pressure might be a misleading surrogate of cardiovascular outcome
  - such unfavorable effects may be unrecognized

Surrogate Endpoints: Potential Pitfalls

- True benefit : risk ratio is often unclear
  - Benefit:

A true, clinically meaningful benefit is not demonstrated in the trial. Rather, it is extrapolated from the observed effect on the surrogate. Therefore, the magnitude of the true benefit may not be certain.

  - Risk:

Shorter, smaller studies using surrogate endpoint may not reveal all risks (i.e. rare, serious adverse effects, or cumulative adverse effects of long-term use). Since patients won't derive meaningful benefit until after long-term use, the magnitude of the corresponding risk is not certain.

(Biomarker)

- A characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.
- Biomarkers are often used in clinical practice to diagnose/stage a disease, or to predict/monitor response to therapy.
- Biomarkers may be utilized in clinical trials to:
  - explore the effects of an investigational drug
  - assess the promise of a drug in early development (e.g. P2)
    - Does the investigational drug exert the expected pharmacologic activity? If so, at what dose?
- Biomarkers cannot establish a clinically meaningful benefit

Patient Reported Outcomes

- Any report of the status of the patient's health that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else.
  - e.g. symptoms, functioning, or a more global assessment of the effect of the disease on health and functioning from the patient's perspective ("health related quality of life")
- Intuitively desirable. A very reasonable goal of therapy would be to make the patient feel better in some way. Sometimes the benefit of a drug may only be detected or described by the patient.
- Current standards for PRO instrument development, validation, and application in clinical trials reflect increasing sophistication in the field.*

*"Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims" (December, 2009)

PRO Challenges

- Development / Validation
  - Does the instrument measure what you want it to measure?
  - Does it measure what is important to the patient?
  - Does it do both of these in the specific population to be studied?
  - If there are multiple concepts / domains being measured (e.g. HRQOL instrument), do they overlap? Are they weighted appropriately?
  - Is the instrument reliable? (e.g. stable in stable patients)
  - Is the instrument sensitive to baseline differences?
  - Is the instrument sufficiently sensitive to detect change over time?
  - Can it be used in multinational studies (multiple languages, cultures)?

PRO Challenges

- Interpretation
  - Was the instrument used appropriately in the trial?
    - training of staff / patients; timing in relation to other assessments; length of recall required; drugs for which blinding may be incomplete (e.g. due to adverse effects, etc.)
  - What does “an improvement of 22” mean?
    - Results are conveyed in a “score”. It may be difficult to understand the magnitude of benefit that has been demonstrated. (Even beyond the “MCID”). This makes it difficult to balance safety risks against that benefit.
  - Did one item drive the result?
    - An instrument intended to measure the impact of a disease may have many items, representing various symptoms / domains. What can you conclude if a positive result is driven by one or few of the items?
  - How to explain a multi-domain concept (e.g. symptoms, functioning, psychological state, social aspects) in the label, or to a patient?

Composite Endpoints

- A single measure of effect, based on a combination of individual endpoints.
- Particularly useful for drugs that can benefit patients in several ways or if component events are infrequent.
- Examples:
  - “MACE” (major adverse cardiac events): cardiovascular death, non-fatal MI, and non-fatal stroke
  - “Clinical Worsening”: may include categorical decline in functioning, worsening symptoms, addition of a new medication, hospitalization due to the disease, death, etc.
  - HRQOL instruments
- Often analyzed as time to first event, or number of events over the study period.

Composite Endpoints

- Considerations:
  - Each component should itself be clinically meaningful.
  - Ideally, each component would be approximately equally meaningful.
  - The composite should not include individual components for which a treatment effect is not expected.
  - “Success” should not be concluded if driven by a less meaningful component, if there is evidence of a therapeutic disadvantage on a more meaningful component.
  - May complicate communication of the established benefit of a drug.
    - There may often be inadequate evidence to establish a treatment effect on any of the components individually. (If a trial “wins” on MACE, can you conclude the drug improves survival?)

Summary

- For marketing approval, there must be *substantial evidence* (consisting of adequate and well-controlled investigations) *of something that matters*
- The primary endpoint(s) of confirmatory Phase 3 trials should represent (directly or through a validated surrogate) something that matters to a patient.
- There is a pathway for approval based on *something that probably matters* (an incompletely validated surrogate), but this comes with certain commitments. ("Subpart H" "accelerated approval")
- Patient reported outcome (PRO) instruments and composite endpoints may be used to establish a benefit that matters.