

Assessing Endpoints in Oncology Clinical Trials

Steven Hirschfeld, MD PhD

Office of Cellular, Tissue and Gene Therapy
Center for Biologics Evaluation and Research

FDA

Summary

- Drug development is an orderly process designed to minimize risk and determine benefit
- Clinical studies are required to produce the evidence to determine risk and benefit
- Clinical development is most effective with understanding and communication between all involved

Why Evidence?

- Law
- Regulations
- Ethics
- Science
- Business

Laws and Regulations for Drugs

- Applicable law is Food Drug and Cosmetic Act as Amended Title 21 - Chapter 9 - Subchapter V - Part A - Section 355 Subsection 1(A)
 - full reports of investigations which have been made to show whether or not such drug is safe for use and whether such drug is effective in use;
- Applicable regulation is Code of Federal Regulations Title 21 Part 314 Section 314.126
 - The purpose of conducting clinical investigations of a drug is to distinguish the effect of a drug from other influences, such as spontaneous change in the course of the disease, placebo effect, or biased observation.
 - 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.

Separate and Parallel Law and Regulations for Biologics

- Applicable law for Biologics is Public Health Service Act Title 42 - Chapter 6a Subchapter II - Part F - Sec. 262. Subsection (2)(B)
 - The Secretary shall approve a biologics license application on the basis of a demonstration that the biological product that is the subject of the application is safe, pure, and potent;
- Applicable regulation for Biologics is Code of Federal Regulations Title 21 Part 601 Section 601.25
 - Proof of effectiveness shall consist of controlled clinical investigations as defined in Sec. 314.126 of this chapter, unless this requirement is waived on the basis of a showing that it is not reasonably applicable to the biological product or essential to the validity of the investigation, and that an alternative method of investigation is adequate to substantiate effectiveness. (*Latter applies only to products approved prior to 1972*)

Why Clinical Trials?

- Marketing claims should be based on evidence
- Evidence should minimize bias and uncertainty
- Bias is any factor that will influence the outcome or interpretation of the results of a study
- Uncertainty is a measure of the confidence in the results

Goal

- Predict the future- not judge the past

Hierarchy of Outcomes

- Live longer
- Live better
- Show clinical activity
- Show biological activity

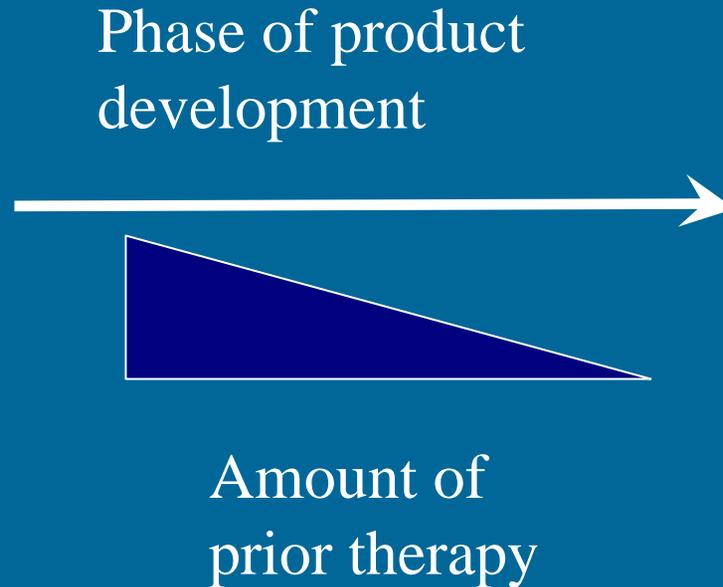
Adequacy of Study Design

- A study that is not likely to yield interpretable results is considered an unethical study because it places participants at risk without the prospect of either direct or generalizable benefit

Section 312.87 Active Monitoring of Conduct and Evaluation of Clinical Trials

- For drugs covered under this section, the Commissioner and other agency officials will monitor the progress of the conduct and evaluation of clinical trials and be involved in facilitating their appropriate progress.

Inverse pattern between product development and prior therapy



Assumptions in clinical development

- Biological activity will translate into clinical benefit
- Activity in treatment experienced patients will translate into activity in treatment naïve patients

However, in any given circumstance these assumptions may not apply

Definition of Study Endpoint

A *study endpoint* is a measure that represents patient benefit or risk for a particular condition. It should:

- have a validated methodology to measure it accurately and reproducibly
- reflect some aspect of the disease process
- change in the same direction as the overall disease process; that is, if the disease progresses the endpoint will change to show the progression and if the overall condition improves the endpoint will change to show the improvement

General Endpoint Characteristics

- Endpoints must be prospectively defined to minimize bias
- The measuring technique must be reproducible and uniform in a patient population
- Serial measurements must be performed consistently with regard to technique and time for all patients at all sites and in all study arms
- Statistical analysis of the endpoints must be prospectively defined
- To be interpretable, must measure a parameter that is clinically relevant or be a validated surrogate

Endpoints with fixed time

- If time is a constant, then the study results are expressed as a rate defined as number of events per time (usually duration of the study)
 - Examples are percentage responders at the end of the study or maximally tolerated dose after one (or more) doses
 - In a comparison, the preferred therapy is usually the one with the greater number of events

Endpoints with variable time

- If time is not constant, but varies, then the results are stated as a length of time or duration
 - Examples include time to progression or survival time or drug half life. These are often expressed as a median of the available data
 - The preferred therapy is usually the one with the longest time

General Types of Endpoints Used in Clinical Oncology

- Survival (usually expressed as a duration)
 - Overall
 - Disease-free
- Time to Progression (duration)
 - Tumor (usually based on imaging results)
 - Onset or worsening of disease related symptoms
- Response (rate)
 - Tumor (usually based on imaging results)
 - Patient Benefit (Palliation, Improvement in symptoms)

General Types of Endpoints Used in Clinical Oncology

- Protection (rate)
 - From adverse events with no decrease in survival
- Reduction in risk of disease (rate)
 - From initial onset in high risk population
 - From recurrence in adjuvant setting

Survival

- Defined as time from randomization to death
- Intent-to-treat analysis (all randomized patients)
- Unambiguous endpoint that is not subject to investigator interpretation or bias from unblinded studies
- Daily assessment
- If primary endpoint then no need to collect other data to demonstrate benefit

Survival-Potential Problems

- Sample size may be large if effect is small
- Follow-up must be adequate with sufficient number of events
- Effect could be due to other factors
- Cross-over design or secondary therapy may blunt or confound effect

Time to Progression/Progression Free Survival

- Defined as time from randomization to documented disease progression for TTP or to disease progression or death for PFS
- Usually smaller sample size and shorter follow-up than survival trials
- Progression might correlate with relief of active symptoms or delay of new or worsening symptoms (difficult to show)

Potential Problems with TTP/PFS

- Usually determined only at fixed intervals with inherent margin of error
- Multiple events, not all of comparable consequence, may be included in definition (for example clinical event, biochemical change, imaging change)
- Imprecision, lack of consistency and potential bias in measurement
- Possibility of investigator bias in unblinded studies
- May not correlate with survival or other clinical benefits, particularly if not symptomatic
- How much improvement constitutes benefit?

Time to Treatment Failure

- Composite of toxicity, disease progression, patient dropout, and miscellaneous factors
- Not used to date for licensing due to difficulties with interpretation

Response Rates

- Effect interpreted as due to agent rather than natural history of disease
- Historically accepted as approval basis for hormonal agents, biologics, and as surrogate for accelerated approval
- Requires precise definition criteria
- Stable disease not part of definition. Implications for cytostatic agents and immunotherapy

Defining Response

The general rules for defining response for solid measurable tumors are:

- the sum of the measurements of all the tumors that can be detected in a single patient must decrease by at least 50% for 2 dimensions or 40% for one dimension
- there can be no new tumor sites (new lesions)
- the decrease must be confirmed by a second set of measurements at least 4 weeks after the initial documentation of response

Further Comments on Response

Response is such a general term that studies sometimes take unconfirmed or arbitrary criteria and define something that is called a response. When a patient meets these criteria, they are scored as having a response although the criteria may never have been demonstrated to predict or reflect patient benefit.

Potential Problems with Response Rate

- Failure to define sites prospectively
- Failure to use same imaging modality
- Failure to identify exact lesion despite same modality used
- Difficult to document response in bone
- Determined only at fixed intervals with inherent margin of error
- Imprecision, lack of consistency and potential bias in measurement

Potential Problems with Response Rate

- Multiple events, not all of comparable consequence, may be included in definition (for example clinical event, biochemical change, imaging change)
- Possibility of investigator bias in unblinded studies
- May not correlate with survival or other clinical benefits, particularly if not symptomatic
- How much improvement constitutes benefit?

Patient Reported Outcomes

- What happens to the patient instead of what happens to the tumor can be a valid outcome
- Examples are
 - decrease in pain
 - improvement in symptoms that limit activity or function
 - delay in the onset of symptoms that limit activity or function
- Valid reproducible measurements are still needed

Patient Reported Outcomes

- Pre-defined disease related problems or symptoms to be measured
- Patient assessed or observed (proxy reporting)
- Prospective
- Hypothesis-driven
- Can directly demonstrate clinical benefit
- Are validated in the target population prior to the study
- Are sensitive to disease specific changes both positively and negatively
- Are contemporary with events-not dependent upon recall

Potential Problems with Patient Reported Outcomes

- Lack of validation in the relevant patient population
- Cultural and literacy differences
- Lack of blinding
- Missing data
- Multiple endpoints
- Challenging analysis
- Lack of sensitivity and correlation with other outcomes
- Time difference between event and recall

Surrogate Endpoints

A *surrogate endpoint* is a substitute for a clinical benefit that is generally more readily measured and may yield a result sooner than assessing the original clinical benefit. A surrogate should have the properties of any endpoint

A surrogate should accurately reflect the clinical benefit it represents

- The surrogate can be measured precisely and reproducibly
- The surrogate changes in proportion to what it represents

Composite Endpoints

- Determined by prospective algorithm
- Analysis may include analysis of each component as well as the composite
- Must include guidelines for interpretation
- If driven by one domain or topic, then claim may be restricted to that domain
- Each component should be of similar clinical significance and potentially observable in each patient

Study Conduct

- Sloppy studies obscure differences.
- If the objective is to prove superiority, a sloppy study is a disadvantage.
- If the objective is to demonstrate similarity (non-inferiority or statistically not worse than), a sloppy study is an advantage

What about safety?

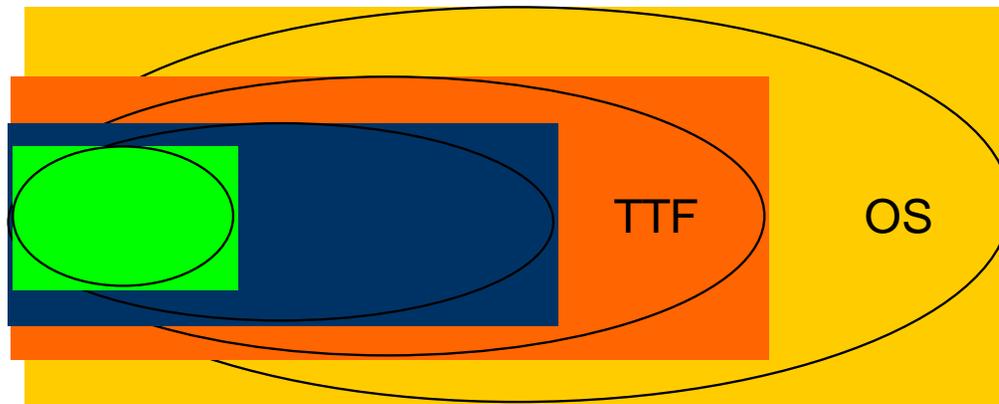
- Society accepts risks for cancer therapy not generally tolerated in other therapeutic areas
- Safety concerns can almost always be managed if a product demonstrates efficacy
- Oncology clinical trials generally use an adverse event scale developed by the National Cancer Institute that must be subsequently mapped onto MEDDRA or another glossary for product labeling

General Types of Analyses Used in Biometrics

- Rate (absolute)
 - Example: Response
- Rate (landmark)
 - Example: 2 year survival
- Time to event
 - Example: Median Progression Free Survival
 - Example: Median overall survival (usually when 70 to 80% of patients have died)

Proportion of Total Enrolled Patient Population Described in Outcome for Variable Time Endpoints

Duration Endpoints (Variable Time)



DFS= Disease Free Survival = Complete Responders

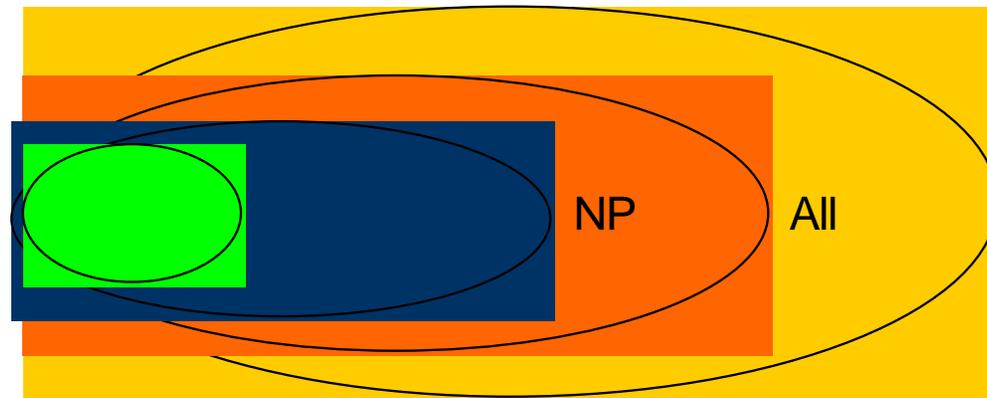
PFS= Progression Free Survival = Non-Progressors

TTF= Time to Treatment Failure= Non-Progressors plus Not Tolerated Toxicity

OS= Overall Survival = All Survivors

Proportion of Total Enrolled Patient Population Described in Outcome for Fixed Time Endpoints

Rate Endpoints (Fixed Time)



CR= Complete Responders

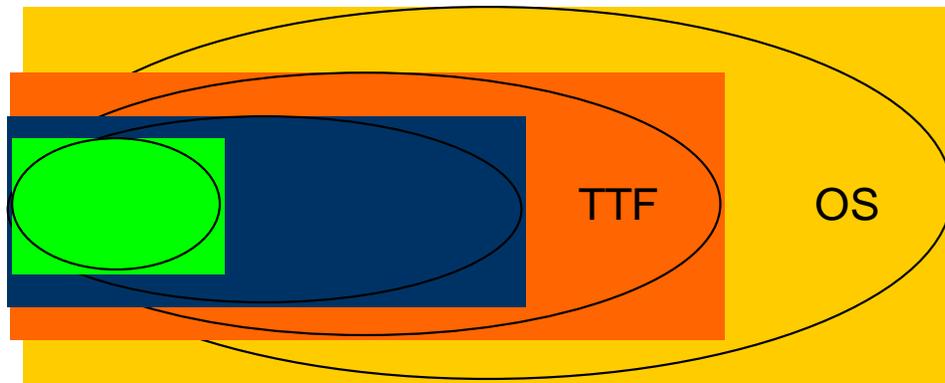
OR= Overall Response

NP= Non-Progressors

All = All patients intended to treat

Proportion of Total Enrolled Patient Population Described in Outcome for Landmark Time Endpoints

Landmark Endpoints (Fixed Time)



DFS= Disease Free Survival = Complete Responders

PFS= Progression Free Survival = Non-Progressors

TTF= Time to Treatment Failure= Non-Progressors plus Not Tolerated Toxicity

OS= Overall Survival = All Survivors

Intent to Treat

To avoid introducing arbitrary criteria as to which patients got the right doses or got enough doses, etc. and thus biasing the results, the most objective measure comes from using all the patients that met the eligibility criteria of the protocol, were registered for the protocol, and were intended to be treated, no matter how much or how little treatment they actually got.

Definition of Efficacy

Efficacy is a term applied to controlled clinical studies that describes the effect of an intervention on the endpoints of interest. For product licensing, efficacy should represent patient benefit. If there were no controlled studies, but patient benefit was recorded, then an intervention would be called effective. When a new treatment is approved, the assumption is that efficacy will predict effectiveness.

Which patients should be analyzed?

For efficacy, all patients in the intent to treat population. The moment that patients are excluded, any randomization is invalid and bias appears in the analysis regardless of study design.

For safety assessment or risk calculation, all patients that were exposed to the intervention.

Is there value to compare Responders to Non-responders?

Occasionally results are reported so that patients that were scored as responding to an intervention are shown to have a better outcome than those that did not respond. This type of analysis defines new groups after the randomization and introduces a new bias, therefore is not a valid comparison. In addition, it is impossible to distinguish between whether the intervention had a benefit or just selected patients who were likely to have a better outcome. In that case the intervention would not be a treatment, but a screening test.

Analyses

- P value does not necessarily translate into patient benefit
- Post hoc analyses are always suspect and should be viewed as hypothesis generating
- Regulatory determinations are based on the totality of evidence

What about endpoints not in the study protocol?

In no circumstances should an endpoint be added to the analysis after the study is completed. It is not possible and not valid to interpret such an endpoint. If an apparent difference appears between the groups for an endpoint not prospectively intended to be analyzed, a new study should be designed using that endpoint as a primary endpoint.

Criteria for licensing a claim

- Demonstration of efficacy with acceptable safety in adequate and well-controlled studies
- Ability to generate product labeling that
 - Defines an appropriate patient population for treatment with the product
 - Provides adequate information to enable safe and effective use of the product

Criteria for licensing a claim

- Examples of specific criteria are:
 - Improved survival or
 - Palliation of symptoms with no decrease in survival or
 - Protection against adverse events with no decrease in survival or
 - Reduction in risk

Benefit that has supported licensure

- Survival--no subjective interpretation
- Prolongation in time to recurrence or disease-free survival (adjuvant trials)
- Prolongation in time to progression or progression free survival
- Palliation (usually with objective response)

Observations

- Multiple pathways to marketing license
- Most therapeutic development begins in relapsed/refractory setting
- Previous therapy must be clinically meaningful
- Patient benefit is either established or implied
- Controlled trials are most informative and required by regulation

FDA Initiatives

- Special Protocol Assessment program

<http://www.fda.gov/cder/guidance/3764fnl.htm>

- FDA-Sponsor meetings with patient representatives

<http://www.fda.gov/oashi/cancer/pconback.html>

- Electronic IND submissions for biologics

<http://www.fda.gov/cber/gdlns/elecgenrev1.htm>

- In partnership with professional organizations
public discussion of clinical trial endpoints

Additional Resources

- Oncology Tools Website
 - <http://www.fda.gov/cder/cancer>

Summary

- Drug development is an orderly process designed to minimize risk and determine benefit
- Clinical studies are required to produce the evidence to determine risk and benefit
- Clinical development is most effective with understanding and communication between all involved