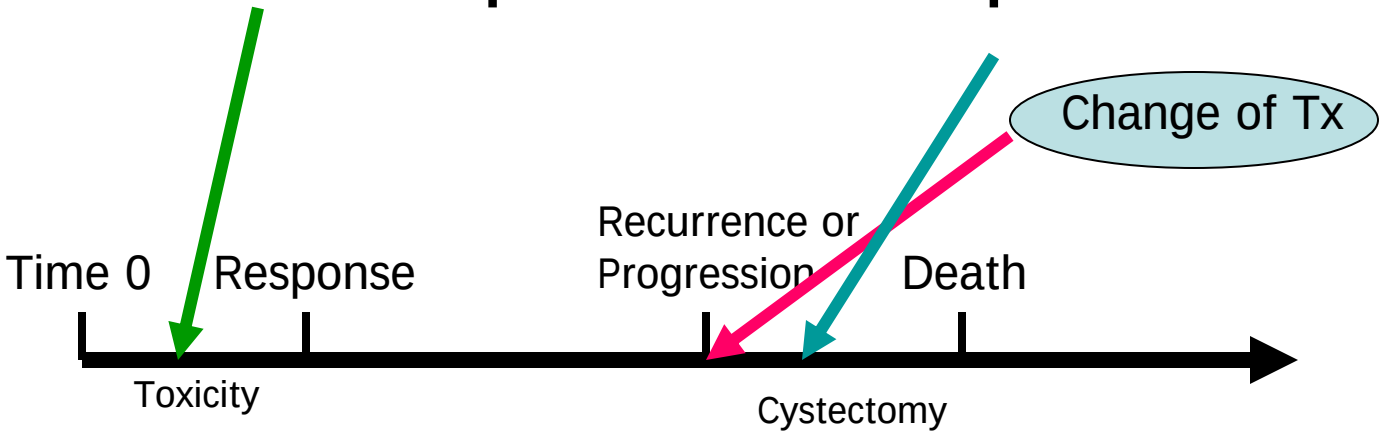
Statistical Considerations in Designing Randomized Clinical Trials

R. Sridhara, Ph.D.
Division Director
Office of Biostatistics, CDER

Design of Clinical Trials

- Sample Size: Using the primary end-point of interest, the required sample size is calculated based on -
  - * Smallest difference in treatment efficacy which is important enough to detect
  - * Acceptable level of false positive result
  - * Acceptable level of false negative result

Sequential Endpoints



Endpoints that can be considered: Response rate, Time to recurrence, Recurrence-free survival, Disease-free survival, Event-free survival, Overall Survival

Time to Event Endpoints

- Efficacy claims based on randomized clinical trials
- EVENT of interest must be clearly pre-defined.
- Not all patients enter the study at the same time.
- Time to Event analysis – some patients may not experience the event at the end of study period/ observation period or at the time of analysis.
- Thus such observations with no documented event are censored.

Censored Observations

- Right Censoring: Event of interest not occurred at the time of analysis – actual event time greater than the amount currently observed. (Almost always when we refer to censoring we mean right censoring)
- Left Censoring: Disease state (event) is observed but do not know when it began. Example: Time from sero-positivity to AIDS
- Interval Censoring: Event occurred between two time points. Example: Recurrence, Progression
- In all our analysis methods we assume that censoring mechanism is independent of the outcome.

Definition: Overall Survival (OS)

Event = Death

Censor: at last date when patient was known to be alive if patient is lost to follow up or is still alive.

Time = Death/censor date – randomization date

Definitions: Event-free survival (EFS)

Event = Failure to achieve CR or Disease Progression/Recurrence or Death (composite endpoint) – whichever occurs first

Censor: at last date of evaluation for recurrence if patient is lost to follow up or is alive and has no documented recurrence

Time = Event/Censor date – randomization date

- If many deaths prior to recurrence – biased estimates

Event Rate – Not Acceptable

Event Rate =

# of events / # of patients randomized

- Ignores Time – could be at 6 months, 1 year, 2 years or 3 years,… Also events occurring at 3 months no different from those occurring at 1 yr.
- Ignores Censored Observations – For example: Some patients may be lost to follow-up, changed treatment, or too early in the trial to have observed an event. Count them as event or not?

Clinical Trials to Demonstrate Efficacy

- Superiority Claim
  - Direct evidence of benefit
- Non-inferiority Claim
  - Indirect evidence of benefit
  - Interpretation could be misleading

Non-inferiority Claim Implies:

1. The effect of T and C are close
2. T and C may be both equally beneficial or

   T and C may be both equally not beneficial

T = Treatment; C = Control

Types of Hypotheses

- Superiority
  = Drug T better than Placebo (P) or
  T better than Active Control (C),
  – C may or may not have known efficacy
- Non-inferiority (NI)
  = T not much less effective than C
  ≠ 'Was Not Different' or 'Similar'
  – C (example: BCG) must have known efficacy and able to quantify the advantage of C over P

Assumptions for NI consideration

- C has an effect (compared to placebo, P)
  - Therefore, C can not be another experimental therapy in NI trial
- Can reliably estimate C effect size

  * C effect size compared to P
- Constancy Assumption

Constancy Assumption

- Control (C) effect in the future trial will be same as the observed effect in the historical trial
  - Population
  - Endpoint measured
  - Dosing Schedule
  - Dose
  - Supportive care, etc

Non-inferiority or Superiority

- Challenges with Non-inferiority Trial
  - How well we know the effect of the control
  - How much of the control effect can we afford to give up
    - At least ‘Δ’% retention of the control effect necessary to make sure that the new treatment is significantly better than placebo

Sample Size Depends On -

- How good the historical data is:
  - One randomized study – large confidence interval – more uncertainty about the effect size
  - Many randomized studies – smaller confidence interval – unless there were fundamental differences in the conduct of the studies: population, dose, etc.

Summary - 1

- Time to event should be the endpoint and not event rate
- Superiority trials provide direct evidence of benefit
  - New drug can be compared to placebo or control (control efficacy established or not)
- Non-inferiority trials provide indirect evidence of benefit
  - New drug must be compared to established control
  - Interpretation can be misleading

Summary - 2

- For NI trial consideration, active control effect must be well characterized
  - Able to estimate effect size
  - Control effect is same now as before
- NI trials are generally large
  - Sample sizes for a NI trial is dependent on the estimate of control effect, population, % retention and the alternative hypothesis
- In considering NI trial, any potential loss of efficacy must be weighed against risk-benefit ratio
- Failed superiority does not imply non-inferiority