2024 Summer Institute In Statistics for Clinical & Epidemiological Research

Module 3:

Design, Conduct, and Analysis of Randomized Clinical Trials with Time to Event Primary Endpoints

Lecture 2:

Clinical Estimands

Scott S. Emerson, M.D., Ph.D.
Professor Emeritus of Biostatistics
University of Washington

1

1

RCT Settings Using Time to Event Outcomes

- Overall goal: Drug discovery
- **Estimands**
  - **Clinical**
  - RCT
  - ICH E9 (R1) strategies for intercurrent events
- Why an “event”? Why ”time to event”?
- Why incomplete observations: Informative vs noninformative?
  - Administrative censoring
  - Competing risks
  - Intercurrent events and protopathic events
  - Loss to follow-up
- How to define “tends to be”?
  - Choice of summary measure

2

2

Clinical Estimands

Scientific Goals

Where am I going?

The statistical jargon "estimand" made it into the general RCT community following the NRC monograph on handling missing data in RCT.

Owing to the common setting of incomplete (censored) time to event data, many investigators have come to erroneously believe that censored data methods will handle all missing data.

(Later, we will consider the processes by which we might be able to estimate these in a RCT.)

3

3

2010 NRC Monograph

- *The Prevention and Treatment of Missing Data in Clinical Trials*
  - Oversight Committee: Experts in missing data methodology and clinical trial methodology

Roderick Little, Chair	
Ralph D'Agostino	Susan Murphy
Kay Dickersin	James Neaton
Scott Emerson	Andrea Rotnitzky
John Farrar	Daniel Scharfstein
Constantine Frangakis	Weichung (Joe) Shih
Joseph Hogan	Jay Siegel
Geert Molenberghs	Hal Stern

4

4

Workshops: What I Learned

- Mission 0a: Consolidation of Clinical Trial Terminology
  - Safety, efficacy, effectiveness
    - What is the estimand?
  - Definition of treatment
    - Treatment versus strategy
  - Study design
    - Standard cohort; placebo vs active run-in
  - Timeframe for primary endpoint
    - Event time, study time, calendar time
  - Multiple endpoints
    - Composite vs co-primary vs primary & secondary
  - Study termination
    - Completion of protocol, stop intervention, consent withdrawn
  - Analysis populations
    - ITT, mITT, per-protocol, safety

5

5

Workshops: What I Learned

- Mission 0a: Consolidation of Clinical Trial Terminology
  - **Safety, efficacy, effectiveness**
    - **What is the estimand?**
  - **Definition of treatment**
    - **Treatment versus strategy**
  - Study design
    - Standard cohort; placebo vs active run-in
  - **Timeframe for primary endpoint**
    - **Event time, study time, calendar time**
  - **Multiple endpoints**
    - **Composite vs co-primary vs primary & secondary**
  - Study termination
    - Completion of protocol, stop intervention, consent withdrawn
  - Analysis populations
    - ITT, mITT, per-protocol, safety

6

6

Workshops: What I Learned

- Mission 0b: Consolidation of Missing Data Terminology
  - Mechanisms generating missing data
    - Toxicity, efficacy (or lack), no longer relevant
    - Sloppy data capture, loss to follow-up, withdrawn consent
  - Statistical definition of missing data mechanisms
    - MCAR, MAR, MNAR
  - Statistical impact of missing data mechanisms
    - Ignorable/non-ignorable
  - Statistical methods
    - Direct imputation (LOCF, BOCF), MMRM, multiple imputation, pattern mixture, weighting
  - Types of sensitivity analyses
    - About assumptions of MCAR, MAR, MNAR
    - About assumptions of analytic models

7

7

Workshops: What I Learned

- Mission 0b: Consolidation of Missing Data Terminology
  - **Mechanisms generating missing data**
    - **Toxicity, efficacy (or lack), no longer relevant**
    - Sloppy data capture, loss to follow-up, withdrawn consent
  - **Statistical definition of missing data mechanisms**
    - **MCAR, MAR, MNAR**
  - Statistical impact of missing data mechanisms
    - Ignorable/non-ignorable
  - Statistical methods
    - Direct imputation (LOCF, BOCF), MMRM, multiple imputation, pattern mixture, weighting
  - **Types of sensitivity analyses**
    - **About assumptions of MCAR, MAR, MNAR**
    - About assumptions of analytic models

8

8

Common Problems (Report)

- Missing data due to discontinuation of treatment
  - Adverse events vs lack of efficacy vs efficacy
  - Specified by protocol vs perception of subjects or investigators
    - Relevance of data *vis a vis* health status, rescue therapies
- Outcomes undefined or unmeasurable for some patients
  - Counterfactual estimands (e.g., QoL after death)
  - Competing risks (e.g., renal function after transplant)
- Missing data because of attrition in the course of the study
  - Missed visits, loss to follow-up, withdrawal of consent
- Missing data in composite outcomes
- Missing data due to death

9

9

Basic Principles

- The missingness must hide a potentially useful value
- The estimand must be scientifically (clinically) relevant
- Reasons for missing data must be documented fully
- Trial designers should decide on primary assumptions about missing data mechanisms
  - Necessarily subjective and untestable
- A statistically valid analysis under those assumptions should consider both consistency and variability of estimates
- The robustness of the conclusions to the untestable assumptions should be investigated

10

10

Classification of Missing Data

- Missing completely at random (MCAR)
  - Indicator of missingness does not depend upon any measured data
    - Sometimes confused with ignorability
- Missing at random (MAR)
  - Within groups defined by some observed data, the data is missing completely at random
  - Information about missing data can be borrowed from data that is available
- Missing not at random (MNAR)
  - Even after conditioning on all observed data, the subjects missing data would have outcomes distributed differently than those for subjects with observed data
- **Nothing in your data can distinguish MNAR from MAR**

11

11

Impact of Missing Data on Analyses

- Ignorable
  - Weak: Analyzing complete cases in the planned analyses provides unbiased estimates of the desired estimand
    - MCAR
    - MAR if we were going to adjust anyway
  - Strong: Just as precisely?
- Nonignorable
  - Failure to account for missingness results in biased estimation of the desired estimand

12

12

Prevention of Missing Data

- The most important issue
  - Through diligence and hard work, experienced RCT investigators have managed to greatly reduce prevalence of missing data
  - It can be done
- Recommendations 2 – 8
  - Design, conduct, analysis
  - Investigators, subjects
- But not discussed further in this course, except as they pertain to proper definition of scientific estimands

13

13

Recommendations for Analysis

- #9: Statistical methods for handling missing data should be specified by clinical trial sponsors in study protocols.
  - Assumptions should be understood by clinicians.
- #11: Parametric models in general, and random effects models in particular, should be used with caution, with all their assumptions clearly spelled out and justified.
- #12: The primary analysis of the data from a RCT should account for the uncertainty attributable to missing data so that type I error rates and associated CI are valid
- #13: Weighted generalized estimating equations methods should be more widely used in settings when missing at random can be well justified and a stable weight model can be determined

14

14

Recommendation # 15

- Sensitivity analyses should be part of the primary reporting of findings from clinical trials. Examining sensitivity to the assumptions about the missing data mechanism should be a mandatory component of reporting.

15

15

ICH E-9 Addendum

- Glass half full or half empty?
- Stresses importance of clearly defining the estimand
  - Does not distinguish between clinical, regulatory, scientific, statistical issues
- Inadequate stress on the need for unbiased ascertainment of clinically important estimands
  - Naïve readers might walk away believing the Addendum sanctions estimands based on something other than randomization
  - Does not address one of the major issues with per protocol analyses
    - Protopathic and indication bias

16

ICH E9 R(1)

- **Estimand:**
  - A precise description of the treatment effect reflecting the clinical question posed by the trial objective. It summarises at a population-level what the outcomes would be in the same patients under different treatment conditions being compared.

17

17

Causal Clinical Estimands

- Compare outcome to counterfactual
  - What happens if on Experimental vs if on Control
- Individual level: Never truly observable
  - ”No man ever steps in the same river twice, for it is not the same river, and he is not the same man” -- *Heraclitus (the Obscure)*
- Population level: Randomized clinical trials
  - Based on (at least conceptual) RCT
  - Try to define a population as precisely relevant to individuals as possible
  - Estimate population average effect
- (Recall, major goal of statistics is to try to stop overfitting of data)

18

Ultimate Goals

- The ultimate goals of basic science, clinical science, and regulatory agencies are somewhat different
- These goals are manifest in all aspects of a treatment indication

19

19

Definition of Disease

- Basic science: Ideal
  - Defined by cause of disease
- Clinical science: Moving target
  - Putative cause vs constellation of symptoms, signs
  - Sometimes reflects response to prior therapy
    - Second line chemotherapy, MRSA, etc.
  - Ultimately refined according to effective therapies
- Regulatory
  - Reproducible definition

20

20

Definition of Population

- Basic science
  - Restrictions based on contraindications of treatment
- Clinical science
  - Considers perceived risks
  - Considers alternative therapies
- Regulatory
  - Legal criteria for efficacy
  - Legal criteria for safety of public
  - Reproducibility
    - E.g., diagnostic criteria for genetics

21

21

Definition of Treatment

- Basic science
  - Effect of precisely defined formulation, dose, administration, frequency, duration, concomitant treatments
- Clinical science: Treatment strategy encompassing some of
  - Modifications of dose, frequency, etc.
  - Prophylactic or concomitant control of adverse treatment effects
  - Rescue and follow-on therapies
- Regulatory
  - Safety margins
  - Reproducibility of treatment definitions

22

22

Treatment: Points Meriting Extra Emphasis

- In a population receiving the same treatment strategy, different patients may at different times receive
  - Different drugs: prophylaxis, rescue, etc.
  - Different formulations: oral, IV, etc.
  - Different doses: titration, escalation, reductions due to AEs
  - Different frequencies: delays, drug holidays due to AEs
  - Different duration: response, lack of response, intolerable AEs
- It is important that both investigators and patients recognize that such differences are all anticipated variations within the same treatment strategy
- Looking ahead: None of these differences should affect our assessment of patient outcomes

23

Definition of Outcomes

- Basic science
  - Intermediate endpoints along causal pathway
- Clinical endpoint
  - Measurable (ethically) for every subject (e.g., anticipate deaths)
  - Long term: clinical benefit
  - Short term: until next treatment decision point
- Regulatory
  - Concordance with public health benefit
  - Concordance with clinical practice
    - Perceived clinical goal (e.g., HTN)

24

24

Scientific Questions of Interest

- We can first consider the types of questions we might want to know the answer to
  - Safety & efficacy vs effectiveness
    - Single endpoints, multiple endpoints, composite endpoints
  - Population defined by treatment compliance
- Conceptually, these can be discussed in a single arm interventional trial

25

25

"Competing Risks"

- Incidence of one event precludes observation of another
  - Time to event analyses: Cause specific mortality
  - All analyses: Withdrawal of consent, loss to follow-up
  - Depending on estimand: Noncompliance, death
- Possible solutions
  - Most important endpoint
    - E.g., overall survival
  - Composite endpoints
    - Progression free survival
    - Quality adjusted life years
    - Major cardiovascular adverse events (MACE)
    - Ventilator free days alive during first 28 days

26

26

Composite Endpoints: Issues

- Clinical relevance of time to first event
- Need to avoid combining endpoints of markedly different clinical importance
  - Death
  - Progression
  - Termination of study drug
- Composites involving invasive procedures require special consideration
  - E.g., liver biopsies may be too risky in some patients
- Regulatory issue
  - How to write an indication for a nonstandard composite endpoint

27

27

Scientific Estimands

- Efficacy of treatment
  1. What is impact among patients who follow protocol?
  2. What is impact among patients who could follow protocol?
  3. What is impact among patients who start treatment?
- Safety of treatment
  1. What is impact among patients who follow protocol?
  2. What is impact among patients who could follow protocol?
  3. What is impact among patients who start treatment?
- Effectiveness of treatment
  - What is impact among patients who would knowingly start treatment?

28

28

Scientific Efficacy Estimand #1

- What is impact among patients who follow protocol?
  - No matter what: An interesting basic science question
    - But not rigorously causal
  - Clinically may be used to explore mechanism of action
    - Sometimes look for stronger effect in per protocol than ITT
- Patients who do not follow protocol are irrelevant
  - Patients who do not follow directions
  - Patients who have intolerable adverse reactions
    - Perhaps "intolerable" only because uncertain of efficacy, or
    - Perhaps leading to serious consequences with continued therapy
  - Patients with real or perceived lack of efficacy
    - Early clinical course is discouraging, or
    - Definitive progression to serious condition prior to primary endpoint
  - Development of contraindication to treatment (e.g., pregnancy)
  - Patients with early evidence of cure

29

29

Scientific Efficacy Estimand #2

- What is impact among patients who could follow protocol?
  - No matter what: A relevant basic science question
  - Depending on safety: Possibly relevant to clinical science
  - Requires estimating outcomes among noncompliant patients
- Some patients who do not follow protocol are irrelevant
  - ~~Patients who do not follow directions~~
  - Patients who have intolerable adverse reactions
    - ~~Perhaps "intolerable" only because uncertain of efficacy,~~ or
    - Leading to serious consequences with continued therapy
  - Patients with real ~~or perceived~~ lack of efficacy
    - ~~Early clinical course is discouraging, or~~
    - Definitive progression to serious condition prior to primary endpoint
  - Development of contraindication to treatment (e.g., pregnancy)
  - ~~Patients with early evidence of cure~~

30

30

Scientific Efficacy Estimand #3

- What is impact among patients who start protocol?
  - No matter what: A relevant basic science question
  - Highly relevant to clinical science
    - But does presume no change in behavior after knowing efficacy
  - No need to estimate outcomes among noncompliant patients
    - We use the measurements made while off treatment
- All patients' data is relevant
  - Hence need to collect efficacy data (in an unbiased fashion) following stopping therapy

31

31

Scientific Safety Estimand #1

- What is impact among patients who follow protocol?
  - No matter what: An interesting basic science question
  - Clinically, may be used to estimate dose response
    - But not rigorously causal
  - Avoids inflating denominator during periods of no exposure
- Patients who do not follow protocol are irrelevant
  - Patients who do not follow directions
  - Patients who have intolerable adverse reactions
    - Perhaps "intolerable" only because uncertain of efficacy, or
    - Perhaps leading to serious consequences with continued therapy
  - Patients with real or perceived lack of efficacy
    - Early clinical course is discouraging, or
    - Definitive progression to serious condition prior to primary endpoint
  - Development of contraindication to treatment (e.g., pregnancy)
  - Patients with early evidence of cure

32

32

Scientific Safety Estimand #1a

- What is impact among patients while they follow protocol?
  - No matter what: An interesting basic science question
  - Clinically, may be used to estimate dose response
    - But not rigorously causal
  - Avoids inflating denominator during periods of no exposure
- Patients who do not follow protocol are irrelevant ***(after drug d/c)***
  - Patients who do not follow directions
  - ~~Patients who have intolerable adverse reactions~~ ***(already counted)***
    - ~~Perhaps "intolerable" only because uncertain of efficacy,~~ or
    - ~~Perhaps leading to serious consequences with continued~~ therapy
  - Patients with real or perceived lack of efficacy
    - Early clinical course is discouraging, or
    - Definitive progression to serious condition prior to primary endpoint
  - Development of contraindication to treatment (e.g., pregnancy)
  - Patients with early evidence of cure

33

33

Scientific Safety Estimand #2

- What is impact among patients who could follow protocol?
  - No matter what: A relevant basic science question
  - If answerable: Definitely relevant to clinical science
  - Requires estimating outcomes among noncompliant patients
- Some patients who do not follow protocol are irrelevant
  - ~~Patients who do not follow directions~~
  - Patients who have intolerable adverse reactions
    - ~~Perhaps "intolerable" only because uncertain of efficacy,~~ or
    - Leading to serious consequences with continued therapy
  - Patients with real ~~or perceived~~ lack of efficacy
    - ~~Early clinical course is discouraging, or~~
    - Definitive progression to serious condition prior to primary endpoint
  - Development of contraindication to treatment (e.g., pregnancy)
  - ~~Patients with early evidence of cure~~

34

34

Scientific Safety Estimand #3

- What is impact among patients who start protocol?
  - No matter what: A relevant basic science question
  - Highly relevant to clinical science
    - But does presume no change in behavior after knowing efficacy
  - No need to estimate outcomes among noncompliant patients
- All patients' data is relevant
  - Hence need to collect safety data (in an unbiased fashion) following stopping therapy

35

35

Scientific Effectiveness Estimand

- What is impact among patients who would knowingly start treatment?
  - Ideally considers benefit / cost tradeoffs through a therapeutic index
  - No matter what: A relevant basic science question
  - Highly relevant to clinical and public health science
    - But does presume no change in behavior after knowing efficacy
  - No need to estimate outcomes among noncompliant patients
- All patients' data is relevant
  - Hence need to collect all data (in an unbiased fashion) following stopping therapy

36

36

Next Lecture

- Overall goal: Drug discovery
- **Estimands**
  - Clinical
  - **RCT**
  - ICH E9 (R1) strategies for intercurrent events
- Why an “event”? Why ”time to event”?
- Why incomplete observations: Informative vs noninformative?
  - Administrative censoring
  - Competing risks
  - Intercurrent events and protopathic events
  - Loss to follow-up
- How to define “tends to be”?
  - Choice of summary measure

37

37