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 3:

RCT 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

Estimands

RCT Goals

Where am I going?

We now consider how we might be able to accurately and precisely estimate the scientific estimands of interest with a rigorous RCT.

3

3

Added Issues in RCT

- RCT are meant to allow the causal effect of the treatment
  - We truly might be interested in within patient effects
  - But these are never truly measurable in the same place, time
  - We thus consider differences between populations who, through randomization, are otherwise comparable
- As we try to quantify the "Scientific Estimands" we face the problem that missing data might not be on comparable subjects
  - We generally do not randomize patients to missingness
- Whenever possible we want an analysis based on randomization

4

4

MCAR in RCT

- Missing completely at random (MCAR)
  - The indicator of missingness does not depend upon any measured data
  - If MCAR, then ignorable
    - Precision might be gained by special analysis, however
- Possible mechanisms
  - By design
    - Measurements made on random subset of subjects
  - By accident
    - Clerical data loss
    - Meteors killing subjects
- MCAR should be rare by accident
  - Can prove missingness is not MCAR, but can not prove MCAR

5

5

MAR in RCT

- Missing at random (MAR)
  - Within groups defined by some observed data, the data is missing completely at random
  - MAR based on pre-randomization variables might be ignorable
- Possible mechanisms
  - Administrative censoring in longitudinal and time to event data
    - Missingness depends solely on date of accrual
    - No time trends in patient characteristics
  - Selected subsampling (e.g., case-cohort studies)
  - Withdrawal of consent or loss to follow-up?
    - Adverse effects, efficacy or lack of efficacy,etc.
    - Possibly differential across arms in incidence and reasons
- Can not use your data to differentiate MAR from MNAR

6

6

MAR Motivating Example: KM

- Administrative censoring in time to event analysis
  - Subjects accrued to study and followed until time of analysis
  - (Presume no time trends in study accrual)
- Subjects with missing data on time of event
  - "Redistribute to the right"
  - We can borrow information from other subjects in the risk set at time of censoring
  - Under noninformative censoring, a censored subject is equally likely to behave like any of the subjects who were still at risk at not censored at that time

7

7

KM: Imputed Data

- KM estimate is in some sense "imputing" the missing data
- We "impute" a censored observation by substituting any of the survival times from others still at risk at the censoring time
  - Each person at risk is equally likely to be used in the imputation
  - We can thus simulate repeated RCT, substituting a randomly selected individual from the risk set for the censored individual
  - We then average the results of the simulated RCTs
- Note that in the case of KM, we can use a formula to perform the multiple imputation

8

8

MNAR in RCT

- 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
- Possible mechanisms (there are zillions)
  - A sudden change in health status
    - is not reflected in any of the scheduled clinic visits / measurements
    - causes a patient to be lost to follow-up or withdraw consent
  - Protopathic signs cause study withdrawal
    - Adverse events are associated with impending events
  - Depending on the estimand, e.g., cause specific mortality
    - Competing risks share a common frailty or tend toward mutual exclusivity

9

9

Possible RCT Estimand #1

- Average improvement for those initially prescribed drug
  - Corresponds to randomized "intent-to-treat" analysis
- Data on all patients is relevant up to the time of the protocol defined primary endpoint
- Unless there is a problem with measurement safety, there should be no missing data from the definition of the estimand

10

10

Possible RCT Estimand #2

- Average improvement for tolerators / compliers
  - An efficacy outcome
  - Safety of trying drug would need to be assessed in another way
- This could be assessed in an RCT using randomized withdrawal or an experimental treatment run-in followed by washout
- Such would eliminate subjects who
  - cannot tolerate due to AEs
  - cannot tolerate due to perception of lack of efficacy
  - are poor compliers
- There are difficulties that need to be considered
  - ”Rebound” effects after discontinuing a drug (possibly off target?)
  - Tachyphylaxis (but detecting this could be the major goal)

11

11

Possible RCT Estimand #3

- Hypothetical average improvement if everyone tolerated
  - This is not directly observable in all cases
  - Requires some sort of modeling of subjects stopping study treatment
    - Models based on MAR, MNAR – unlikely to be MCAR
  - This counterfactual is of very limited clinical importance unless we imagine ancillary treatments that mitigate lack of tolerance
- This could be partially assessed in a RCT with extraordinary incentive
  - Perhaps would handle mild toxicity and mild lack of efficacy
  - Could not be addressed for all cases of stopping study drug
    - Need to avoid coercive incentives
- **(I am very reluctant to use this as a primary estimand)**

12

12

Possible RCT Estimand #4

- Average AUC improvement during adherence
  - Measure efficacy outcome only while adherent
  - Integrate area under the curve
  - Does not require efficacy data following stopping treatment
  - Presumably of interest as a ”hypothetical” strategy
- Incorporates adherence as the timeframe of interest, with both longer adherence and better efficacy reflected in the magnitude of the effect
  - Might in some sense equate two treatments
    - one having low dropout, with mild efficacy benefit
    - one having high dropout, but high efficacy benefit
- This can be addressed in a RCT, providing comfortable with the composite adherence-efficacy endpoint
  - **(I always reject use of this as a primary estimand)**

13

13

Possible RCT Estimand #5

- Average improvement during adherence
  - Incorporates adherence as the timeframe of interest, but length of adherence is averaged out
  - No need for efficacy data after stopping treatment
- This approach would equate two treatments in which
  - one has high efficacy during a short phase of tolerability
  - other has high efficacy during a long period of tolerability
- This approach would prefer a treatment that provided a short burst of high efficacy prior to lack of tolerance over a more tolerable treatment that provided long term moderate effect
- This can be addressed in an RCT if comfortable with the timeframe of measurement
  - **(I do not regard this as a scientifically rigorous measure of efficacy)**

14

14

Assessing Effectiveness

- For all but the first estimand, safety must be assessed separately
- Need to consider safety in the general population, including non-tolerators
  - Short- and long-term AEs from short term exposure
  - Harm from delay of starting efficacious treatment

15

15

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

16

16