Introduction to Clinical Trials - Day 2

Session 7 - Data Management in Clinical Trials

Presented July 26, 2016

Susanne J. May
Department of Biostatistics
University of Washington

Daniel L. Gillen
Department of Statistics
University of California, Irvine

Planning for Data Collection

Paul Dickson, in The Official Rules, Delacorte Press, 1978, gives this:

Stamp's Statistical Probability "The government [is] extremely fond of amassing great quantities of statistics. These are raised to the nth degree, the cube roots are extracted, and the results are arranged into elaborate and impressive displays. What must be kept ever in mind, however, is that in every case, the figures are first put down by a village watchman, and he puts down anything he damn well pleases."

(Attributed to Sir Josiah Stamp, 1840-1941, H.M. collector of inland revenue.)

Planning for Data Collection

Ultimate goal of an RCT

- The goal of a RCT is to find effective treatment indications
- At the conclusion, this will require reporting the experiment

1. Overall goal
2. Specific aims
3. Materials and Methods
   - Patients, dosing, adherence to monitoring
4. Results
   - Disposition, compliance, adverse events, outcomes
5. Conclusions

Planning for Data Collection

Role of Data : Overall goal and specific aims

- Goal / aims ideally determined prior to start of study
- BUT, the question actually answered is specific to
  - the subjects actually sampled
  - the methods actually used
  - the data actually gathered
  - the analysis actually performed
- Generalization of results depends on all of the above

Role of Data : Materials

- Eligibility criteria are usually broad
- Need to describe the population actually sampled
- Need to describe how the sample might differ from the ultimate target population

Role of Data : Materials

Conceptual framework

- Population of patients with disease
  - Definition of disease by cause vs signs / symptoms
- Subpopulation with disease targeted by intervention
  - “Disease" truly defined by treatment?
- Subpopulation eligible for study accrual
  - Restricted due to general clinical trial setting
- Eligible patients from which sampled
  - Restricted due to specific clinical trial (location, time)
- Study sample
  - Restricted due to willingness to participate
- Analysis sample
  - Data collection

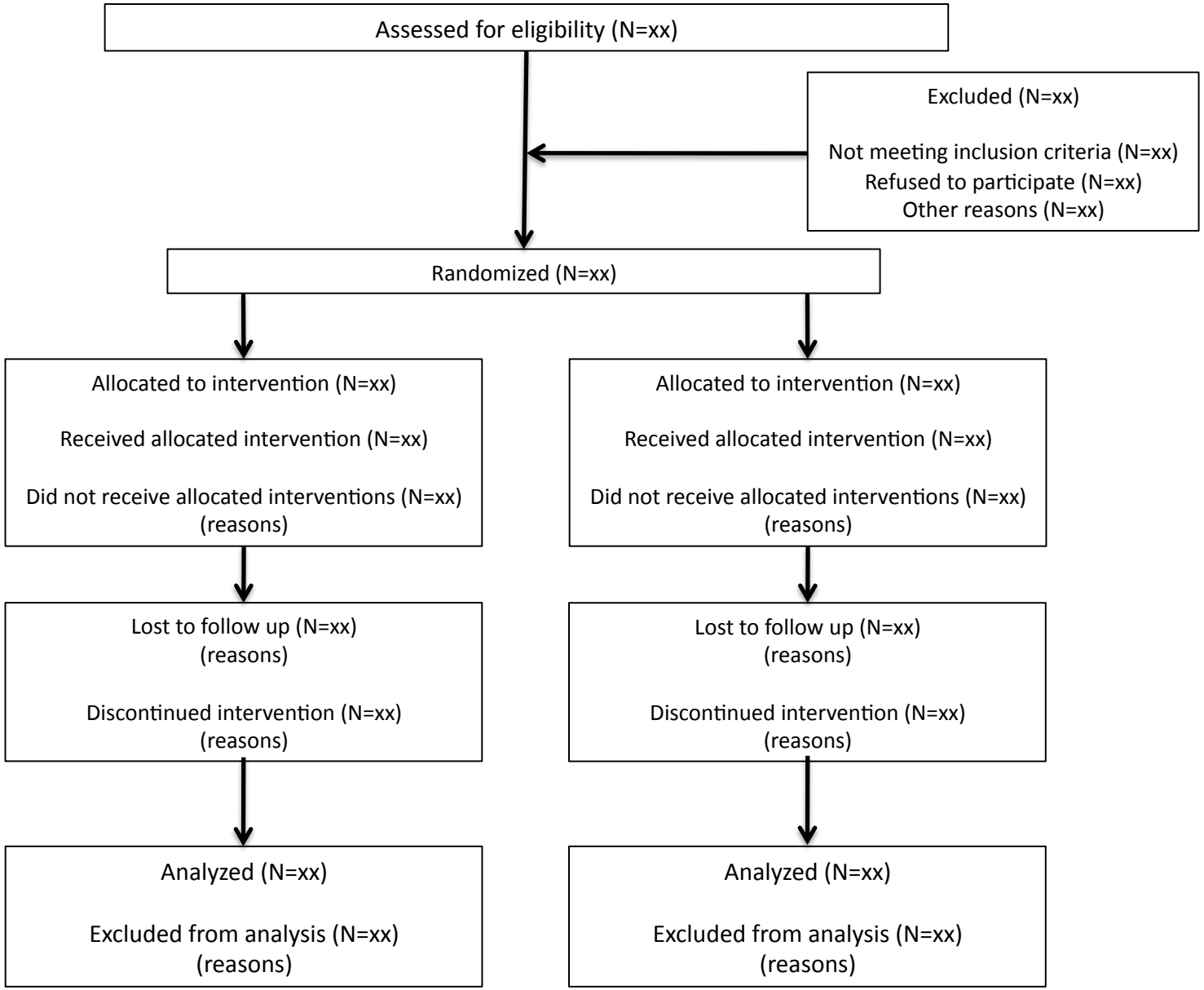
Role of Data : Materials

Generalizability

- CONSORT: Consolidated Standards of Reporting Trials
- Evidence based, minimum standards
- Report flow of patients from screening to collection of primary outcomes
  - Screened
  - Enrolled
  - Randomized
  - Completed

Role of Data : Materials

CONSORT Diagram



Role of Data : Materials

Initial Screening Data

- Source of screened patients
- Number screened
- Characteristics (may require consent)
  - Demographics
  - Disease characteristics
- Reasons for ineligibility
  - Inclusion criteria
  - Exclusion criteria
  - No participation
    - Unable to contact
    - Refused participation

Role of Data : Materials

Screening Visit(s) Data

- Consent for screening
- Contact information: Name, address, alternative contacts...
- Demographics: Sex, age, race, ethnicity...
- Disease characteristics: Duration, severity,...
- Prior and ongoing treatments
- Eligibility data
  - Inclusion criteria
  - Exclusion criteria
- Consent for randomization

Role of Data : Materials

Baseline Visit(s) Data

- Characterize patients
  - Severity of disease, concomitant disease...
- Baseline measures of outcomes
  - Concomitant medications
  - Adverse events
  - Efficacy outcomes (eg. initial tumor size for progression)
- Note differing detail needed for screening vs baseline

Role of Data : Materials

Run-in Data (if applicable)

- Placebo: All patients take placebo
  - Washout vs assessing compliance
  - Patients may be blinded to existence of run-in
- Active: All patients take experimental therapy
  - Allows randomized comparison of efficacy in patients actually taking drug
    - Randomized withdrawal of drug (among “responders"?)
    - Usually patients aware of run-in
  - Assess tolerability for AEs
  - Assess compliance

Role of Data : Materials

Randomization Data

- Documentation of eligibility
- Informed consent
- Stratification variable
- Variables needed for determination of dosing
  - Weight, BSA, renal function, severity of disease...
- Time, date of randomization
- Documentation of assigned group (blinded)
  - Cluster?
- Receipt of first treatment: time, date

Role of Data : Results

Treatment Data : Why

- Intention to treat analysis is the standard for efficacy
- Patients are analyzed in assigned group irrespective of their compliance
- Compliance data is an outcome
  - Assess possible AEs
  - Assess possible mechanism for lack of effect
  - Describe realized exposure to treatment
  - Exploratory analyses for dose / response?
- Safety analyses are typically analyzed according to drug exposure
  - AEs / SAEs occurring within 28 days (?) of last dose

Role of Data : Results

Treatment Data : What

- Initial assignment
  - Dose, administration, frequency, duration, ancillary treatments
- Protocol specified modifications
  - Dose reduction / escalation / holidays
  - Date, time, reasons for change (eg. AE, efficacy or lack of efficacy)
- Patient compliance
  - Dose, frequency, duration
  - ntermittent vs permanent change

Role of Data : Results

Treatment Data : How

- Protocol specified modifications
  - Regularly scheduled visits
  - Interim visits
- Patient compliance
  - Patient diaries
  - Pill counts
  - Clocks on container lids
  - Biochemical measures: blood, biopsies

Role of Data : Results

Patient Monitoring Data : Safety

- Protocol defined safety endpoints
  - Clinical events, subclinical laboratory measurements
- Adverse events
  - Review of interim AEs at regular visits
  - Undesirable clinical events that occur during the study
    - Treatment emergent: new or exacerbated
    - Classification (e.g. MEDRA), grade of severity
  - Treatment relatedness (but do not necessarily believe)
- Serious adverse events
  - Fatal, life-threatening, hospitalization or prolongation, birth defects
  - Expedited reporting if unexpected

Role of Data : Results

Patient Monitoring Data : Efficacy

- Protocol defined efficacy endpoints
- Clinical events
  - Create patient symptoms
- Quality of life
- Subclinical events
  - Signs thought to be indicative of clinical risk
  - Protocol specified monitoring schedules of
    - Patient performance (FEV, 6 minute walk, etc.)
    - Blood
    - Tissue biopsies
    - Radiology

Role of Data : Results

Missing Data : Efficacy

- Lack of training: Patients, investigators
  - Off study drug
  - Decline invasive procedures
  - Withdraw consent
- Poor endpoint definition
  - All randomized patients must have defined outcome
    - E.g, Quality of life after death, GFR in dialysis, symptom relief with noncompliance
- Sloppy conduct of RCT
  - Excessive loss of follow-up

Patient Monitoring Data: Compliance

- Patient adherence to measurement of outcomes
  - Clinic visits
    - Timing relative to window
  - Outcome assessments : Efficacy
    - Blood, tissue samples; radiology, special exams
    - Withdrawn consent for invasive procedures?
  - Outcome assessments : Adverse events
    - Periodic reports per protocol
    - Capture of interim SAEs

Role of Data : Results

Patient Monitoring Data: Logistics

- Patient change of address
  - (sometimes schedule phone visits to maintain contact)
- Site compliance with timeliness completeness

Role of Data : Results

End of Study

- Reason for stopping study
  - Completion per protocol
    - May be off study drug but still followed
    - Death
    - Withdrawn consent (Reasons)
- Permission for further follow-up
  - Change of address
- Conjectured treatment assignment?

Data Collection

Sources of Data

- Subject self report
- Proxy for subject
- Clinic staff and study records
  - Standard medical care
  - Protocol specified procedures
- Medical records
- Laboratory, radiology, pathology
  - Local vs central labs
- Adjudication panels
- Public health records
  - Registries
  - National Death Index?

Data Collection Methods

- Forms
  - Abstracted from medical records
    - Indication bias
  - Completed by subject
  - Completed by proxy
  - Administered by study personnel
  - Completed by clinic staff, study personnel
  - Completed by adjudication panels
- Data files
  - E.g., laboratory, Medicare, National Death Index

Data Collection

- Development of forms
  - Administrative information
    - For follow-up, etc.
    - Often text
  - Scientific information
    - Needs to be appropriate for statistical analysis
    - Free text is difficult to analyze
    - Coding of response by person closest to the source

Data Collection

- Development of forms (cont.)
  - Format of forms should facilitate
    - Completion of form
    - Brief as possible
    - Make sure no portions overlooked
    - “skip patterns", two columns, back of page
    - Cover all cases (explicit “does not apply")
  - Data entry
    - Coding on form

Data Collection

- Issues in form development
  - Number of distinct forms
  - Guidance to the subject, clinic staff on form
    - Study specific definitions
    - Indications for study procedures, other forms
    - Convenience versus increased length
  - Manual and training for form completion
  - Forms for subject vs proxy vs administered
  - Translations
  - Pretesting: subject, staff, investigators, statistician
  - Mapping between different versions over time

Planning for Data Management

- Data to be collected: What? Why?
- Methods of collection: Who? Where? When? How?
  - Forms development
- Methods for data storage
  - Development of database
    - Administrative data: often dynamic
    - Scientific data: usually static
- Methods for data entry
  - Distributed versus central

Data Management

Handling of Data

- Collection
- Data entry
- Storage of forms, primary records
- Data verification
- Checking for errors
- Data reporting
- Data analysis
- Final database

Data Entry and Storage

- Data Entry
  - Transcription of data from forms into computerized data base
  - Personnel often clerical staff
    - Little scientific knowledge
  - Minimize data entry errors
    - Screen for impossible values
    - Screen for inconsistencies within form
    - Double entry
- Storage of forms, primary records
  - Subject confidentiality is a major concern
  - Must ensure limited access to confidential information

Data Verification and Error Checking

- Data entry errors
- Data collection errors
  - Audit clinics
  - Compare study data to medical records
- Maintaining an audit trail
  - Changing database versus making corrections in separate files

Data Reporting

- Administrative analyses
  - Accrual rates
  - Timeliness of data collection
  - Completeness of data collection
- Baseline characteristics
- Event rates (combined treatment groups only)

Data Analysis

- The ultimate purpose of collecting the data
- MUCH easier, more generalizable if all the previous stages conducted properly
- Complete record of all analyses should be maintained
  - date of analysis
  - version of data base and software