Clinical Trials in Rare Diseases: Challenges in Design, Analysis, and Interpretation

Michael P. McDermott, Ph.D.
University of Rochester Medical Center
December 6, 2013
Overview

• Clinical trials in rare diseases present several challenges
  – Such trials are more prone to variability and may have power to detect only large treatment effects
  – Importance of study planning is magnified and planning requires more time
  – Critically important to forge collaboration between clinicians and statisticians
Institute of Medicine Report (2001)

• Several recommendations pertaining to the design, analysis, and interpretation of clinical trials that, for reasons that are unavoidable, are constrained to be small
Institute of Medicine Report (2001)

• Recommendations
  – Define the research question
    • Need to help clinicians make therapeutic decisions
  – Tailor the design
  – Clarify methods of reporting of results in clinical trials
    • Research synthesis; clinical context
  – Perform corroborative statistical analyses
    • Uncertainty regarding analysis assumptions
  – Exercise caution in interpretation
    • Extrapolation of study results
  – More research on alternative designs is needed
Clinical Trials in Rare Diseases

• Limited availability of resources
  – Willing trial participants
  – Funding sources

• In this setting, feasibility constraints can lead to compromises in important principles of sound trial design
Some Important Principles of Sound Trial Design

• Precise formulation of a focused research question
  – Prioritization of outcome variables and analyses

• Tailoring of study design to best answer the research question posed
  – Minimization of bias
    • Randomization
    • Blinding
    • Appropriate control group
      – Context of existing treatment
      – Use of placebo/sham treatment
Some Important Principles of Sound Trial Design

• Tailoring of study design to best answer the research question posed
  – Appropriate eligibility criteria
    • Generalizability vs. efficiency
  – Appropriate outcome measures
    • Reliable, valid, responsive, applicable
    • Duration of follow-up
  – Appropriate and feasible sample size
  – Appropriate measures for participant recruitment and retention
    • Frequency and timing of assessments
    • KISS principle
Some Important Principles of Sound Trial Design

• Use of appropriate and efficient statistical methods for data analysis
  – Parametric models (if appropriate, pre-specified)
  – Covariate adjustment
  – Use of longitudinal data

• The importance of these principles becomes magnified in the setting of a rare disease
  – May require more attention to study planning
    • One cannot increase the sample size, so one must increase efficiency in other ways
  – Importance of collaboration between clinicians and statisticians
Outcome Variables

• Continuous
  – Tend to be more responsive
  – Meaningful?
  – Normally distributed?
  – Replicate measures can increase precision

• Time-to-event
  – Example: Disease milestone

• Categorical
  – Tend to be less responsive
Outcome Variables

• Use of longitudinal data
  – Change from baseline to final visit
  – Use of data from all visits
    • Area under the response-time curve
    • Average of responses after a certain time point
    • Slope (rate of change)
  – Choice may depend on expected timing of onset/loss of maximal benefit
  – Choice also depends on the clinical question that is most relevant to address
Issues that Small Trials Are Better Equipped to Address

• Pharmacokinetics
  – Single- and multiple-dose studies
• Maximum tolerated dosage
  – “3 + 3” or continual reassessment designs (or variants)
• Short-term safety
• “Activity” of the treatment
  – Need for appropriate markers of “activity”
• Preliminary efficacy or futility
  – Acceptance of higher error rates (false positive, false negative)
• Selection of a treatment
Selection Designs

• Goal is to select, out of $k$ potential treatments, the one with the best response
  – Randomized, parallel-group trial
  – Requires a much smaller sample size than a trial designed to formally test the null hypothesis of no treatment effects
    • Sample size is chosen to provide high probability of selecting the better treatment given that the treatments differ by a specified magnitude

• The selected treatment may not be superior to placebo
  – To be confirmed in a separate investigation
Other Designs for Small Trials

- Parallel group design
  - Randomized concurrent controls
  - Natural history controls
- Cross-over design
- N-of-1 design
- Adaptive designs
- Sequential designs
Use of Historical Controls

• Advantages
  – Approximately **four-fold fewer subjects required** compared to a two-arm trial with a concurrent control group
  – Recruitment

• Dangers of using historical controls
  – Changes in ancillary care over time
  – Differences in rater behavior
  – Differences in entry criteria
  – Differences in recruitment of subjects
  – Lack of blinding
  – CMT-1a example
Trial of Vitamin C in CMT-1a
Consequences of the Use of Invalid Historical Controls

• Biases tend to favor treatment under study
• Ability to conduct subsequent confirmatory trials can be compromised
  – “Positive” results from trials that do not use a randomized, concurrent control group
• Treatment can be worse than placebo
  – Recent examples: minocycline and lithium in ALS
• A rare disease is no excuse for a poorly designed study
Cross-Over Designs

• A cross-over trial is one in which subjects are given different treatments during different treatment periods, with the object of comparing the various treatments

• Treatments are given in a randomly determined sequence (e.g., A/B vs. B/A)
## Two-Period Cross-Over Design

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Period 1</th>
<th>Washout</th>
<th>Period 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/B</td>
<td>A</td>
<td>_______</td>
<td>B</td>
</tr>
<tr>
<td>B/A</td>
<td>B</td>
<td>_______</td>
<td>A</td>
</tr>
</tbody>
</table>
Cross-Over Designs

• Appropriate for treatments that may offer short-term relief of signs or symptoms, not a cure for the condition
  – Asthma, hypertension, epilepsy, pain, other chronic conditions

• It is assumed that the symptom or condition will return after withdrawal of treatment
  – Not appropriate for acute or rapidly progressive conditions

• Most appropriate for relatively short-term studies
Advantages of Cross-Over Designs

• Profound savings in sample size
  – Within-subject comparisons
• Participants gain access to all treatments under study
  – May enhance recruitment/retention
Disadvantages of Cross-Over Designs

- Not suitable for all conditions
  - JNCL?
- Impact of subject withdrawal
- Importance of blinding is magnified
- Inconvenience to participants
  - Multiple treatment/washout periods
  - Total duration of follow-up
Disadvantages of Cross-Over Designs

• Treatment-by-period interaction
  – Carry-over effect; other causes
  – Assumption of no treatment-by-period interaction is (virtually) untestable
  – Use of washout periods of sufficient duration
N-of-1 Trials

• Performed in multiple pairs of treatment/placebo periods
  – Example: AB BA BA AB . . .
  – Feasibility of multiple treatment periods
  – Same limitations as those for cross-over trials discussed earlier
• Require rapid onset/washout of the treatment and its effects
• Inference for individual patients is limited without having many periods
• A series of N-of-1 trials in different patients can be much more powerful
  – Random effects models can be used to combine information across patients
Adaptive Designs

• Pharmaceutical Research and Manufacturers of America (PhRMA) Working Group (2006):
  – “By adaptive design, we refer to a clinical study design that uses accumulating data to decide how to modify aspects of the study as it continues, without undermining the validity and integrity of the trial.”
  – “In such trials, changes are made ‘by design,’ and not on an ad hoc basis; therefore, adaptation is a design feature aimed to enhance the trial, not a remedy for inadequate planning.”
Adaptive Designs

• Validity
  – Correct statistical inference
    • Control of Type I and Type II errors
    • Minimization of bias
  – Consistency between stages of the trial
  – Low operational bias

• Integrity
  – Results are convincing to a broader scientific community
  – Pre-planned adaptations
  – Maintenance of the blind to interim analysis results
Some Types of Adaptive Designs

Flexible Designs

Adaptive Designs

Unplanned

Studies with Unknown Properties

Planned

Learning Phase

Combined (Seamless) Phase

Confirmatory Phase

Adaptive Dose Response (Toxicity)

Phase I/IIa

Adaptive Dose Response (Efficacy)

Phase IIb/III

Adaptive Randomization

Group Sequential

Sample Size Re-estimation

Combinations of GS and SSR

Kairalla et al. Trials 2012; 13:145
Adaptive Dose Finding

• Traditional approach in Phase II
  – Randomization to a relatively small number of fixed dosages (3-4) and placebo
  – Disadvantages
    • Large “distance” between adjacent dosages
      – Optimal dosage may not be studied
    • Some of the studied dosages may not be useful
      – This may become apparent relatively quickly
    • Accumulating evidence may suggest early stopping for futility or identification of a sufficient dosage to study further
Seamless Phase II/III Designs

- Dosage A
- Dosage B
- Dosage C
- Placebo

Phase II (Learning) → GAP → Phase III (Confirming)
Seamless Phase II/III Designs

Dosage A

Dosage B

Dosage C

Placebo

Stage 1 (Learning)

INTERIM ANALYSIS (Dosage Selection)

Stage 2 (Confirming)
Group Sequential Designs

• Interim analyses of accumulating data
  – Ethical issues
  – Efficiency/cost
  – Consideration of safety, efficacy, and futility
  – Problem of repeated significance testing
    • Increased probability of “false-positive” (efficacy)
    • Increased probability of “false-negative” (futility)
    • $\alpha$- and $\beta$-spending functions
      – Used to define “stopping boundaries”
OBF = O’Brien–Fleming Bounds [12].
E-F = Emerson–Fleming Bounds [26].
Adaptive Designs

• There are many logistical and procedural issues that are introduced by the possibility of adaptation
  – Careful planning; evaluation of feasibility; infrastructure

• Trial integrity should be preserved by minimizing access to information on interim analyses and their results
  – Control of operational bias
Small is Not Big

• In a very rare disease, sacrifices in important areas may have to be considered
  – Early/middle development
  – Error rates (significance level; power)
  – Use of a less meaningful outcome in favor of a more responsive outcome
  – Randomization/blinding/concurrent controls

• One must fully understand the implications of these sacrifices
Conclusions

• Small studies can be very useful if rigorously and carefully planned, executed, and interpreted
• It is important to formulate specific questions that can be answered with a small study
• An appropriate study design has sufficient sample size and proper control of bias to allow meaningful interpretation of the results
  – Although small clinical trials pose important limitations, these issues cannot be ignored
EXTRA SLIDES
A standard “3 + 3” dosage escalation design starting at dosage k. The maximum tolerated dosage (MTD) is usually defined as the highest dosage at which 0 or 1 dose-limiting toxicities (DLTs) are observed in six patients (although some “3 + 3” rules call the highest dosage with two or fewer dose-limiting toxicities in six patients the MTD). If de-escalation occurs at the first dosage level, then the study is discontinued.
Figure 10.4  Simulated CRM dose-finding. Each panel shows the current data, updated model fit (solid curve), and the previous iteration’s model (dotted curve).
Problems with Many Preliminary Studies

• Often, preliminary studies, particularly in rare diseases:
  – Are very small
  – Are uncontrolled
  – Do not address a focused question
  – Do little to enhance decision-making for further study of the intervention and, as a consequence, slow research progress
Potential Adverse Consequences of Small Trials

• Discarding of potentially effective treatments due to lack of statistically significant benefits
  – “Negative” vs. “Inconclusive” studies
  – P-values vs. confidence intervals
• Inappropriate emphasis on informally defined (or undefined) “trends in the right direction” (or lack thereof)
• Illusion of safety
# Role of Confidence Intervals in Trial Interpretation

<table>
<thead>
<tr>
<th>% Difference in Rate of Progression</th>
<th>95% Confidence Interval</th>
<th>P-value</th>
<th>Evidence for Treatment Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>30%</td>
<td>(-20%, 80%)</td>
<td>0.30</td>
<td>Inconclusive</td>
</tr>
<tr>
<td>30%</td>
<td>(20%, 40%)</td>
<td>0.003</td>
<td>Positive</td>
</tr>
<tr>
<td>2%</td>
<td>(-4%, 8%)</td>
<td>0.30</td>
<td>Negative</td>
</tr>
<tr>
<td>2%</td>
<td>(1%, 3%)</td>
<td>0.003</td>
<td>Positive, but not clinically important</td>
</tr>
<tr>
<td>2%</td>
<td>(-30%, 34%)</td>
<td>0.93</td>
<td>???</td>
</tr>
</tbody>
</table>
## Role of Confidence Intervals in Trial Interpretation

<table>
<thead>
<tr>
<th>% Difference in Rate of Progression</th>
<th>95% Confidence Interval</th>
<th>P-value</th>
<th>Evidence for Treatment Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>30%</td>
<td>(-20%, 80%)</td>
<td>0.30</td>
<td>Inconclusive</td>
</tr>
<tr>
<td>30%</td>
<td>(20%, 40%)</td>
<td>0.003</td>
<td>Positive</td>
</tr>
<tr>
<td>2%</td>
<td>(-4%, 8%)</td>
<td>0.30</td>
<td>Negative</td>
</tr>
<tr>
<td>2%</td>
<td>(1%, 3%)</td>
<td>0.003</td>
<td>Positive, but not clinically important</td>
</tr>
<tr>
<td>2%</td>
<td>(-30%, 34%)</td>
<td>0.93</td>
<td>Inconclusive</td>
</tr>
</tbody>
</table>