Missing Data Issues in Regulatory Clinical Trials

Lisa M. LaVange, PhD
Office of Biostatistics
OTS/CDER/US FDA

JPMA Symposium
February 12, 2015
Acknowledgments

The work presented here is in collaboration with Dr. Thomas Permutt, FDA-CDER representative to the ICH E9(R1) Working Group. Dr. Permutt also chairs an internal (within the Office of Biostatistics) working group on missing data. Other members are:

- Dr. Aloka Chakravarty
- Dr. James Hung
- Dr. Steve Wilson
- Dr. Daphne Lin
- Dr. Raji Sridhara

Three manuscripts have been submitted for publication based on the work of this group and are currently in review.
Outline

• Introduction to the Office of Biostatistics
• Background and motivation – The NRC report
• The issues
  – Prevention
  – Estimands
  – Sensitivity analysis
• Examples
• Summary
INTRODUCTION
Office of Biostatistics

US Food and Drug Administration (US FDA)

Center for Medical Products and Tobacco (CMPT)

- Center for Biologics Evaluation and Research (CBER)
- Center for Drug Evaluation and Research (CDER)
- Center for Devices and Radiological Health (CDRH)
- Center for Tobacco Products (CTP)
2014 Metrics

- 182 statisticians and support staff
  - 300+ statisticians FDA-wide
  - Planned growth to 215 (in CDER)
- ~250 statistical reviews for efficacy, safety, and bioequivalence in NDAs/BLAs/ANDAs/SNDAs
- ~150 non-clinical reviews (stability, CMC, analytical similarity) and other clinical (TQT, abuse liability) reviews
- ~1500 IND/protocol reviews (for ~950 unique INDs)
- 60 peer-reviewed publications
- Over 50 external presentations
Office of Biostatistics

• Vision statement:
  The Office of Biostatistics is recognized for excellence in the application and communication of statistical science in drug regulation and development. We play a central role in promoting innovative, science-based, quantitative decision-making throughout the drug development life-cycle.

• Mission statement:
  Provide CDER and other internal and external stakeholders with statistical leadership, expertise, and advice to foster the expeditious development of safe and effective drugs and therapeutic biologics for the American people. Protect the public health by applying statistical approaches for monitoring the effectiveness and safety of marketed drugs and therapeutic biologic products.
Statistical reviews

• Key components of statistical reviews
  – Determine the accuracy of sponsor’s data and the validity of sponsor’s analyses
  – Assess the robustness of sponsor’s results, particularly when key analysis assumptions are difficult to verify
  – Evaluate results in light of evidentiary standards for confirming efficacy and safety
  – Assess risk and benefit in pre- and post-market reviews
  – Interact with advisory committees

• Strive for consistency across medical divisions while also factoring in therapeutic considerations
• Strive for transparency and reproducibility of reviews
2014 OB Regulatory Science Day
Motivation – The NRC Report

BACKGROUND
Motivation

- 2008: US FDA contracted with NRC for advice
  - Explosion of sophisticated methods introduced to manage missing data during analysis
  - Lively debate among academics and practitioners, with little consensus on best approach
  - Strategies developed for samples surveys or observational studies (e.g., nearest neighbor and multiple imputation methods) not necessarily appropriate for randomized controlled trials (RCTs)
  - Research vs. regulatory perspective: find best model post hoc vs. pre-specification
  - Trialists and investigators needed a wake-up call, i.e., can’t fix everything in the data analysis phase
NRC Report

Published in 2010 by the National Research Council of the National Academies of Science

Panel chaired by Prof. Rod Little
NRC Report

• Published in 2010
  – Heavy emphasis on designing and conducting studies to avoid or prevent missing data
  – Surveyed methods, but did not recommend any one method or class of analysis methods
  – Strongly advised against the use of single-valued imputation methods (last observation carried forward, baseline observation carried forward, etc.)
  – Advocated conducting sensitivity analyses to assess the impact of missing data on the study’s results
Aftermath

- Issues encountered in regulatory reviews since publication of the NRC report:
  - Some increase in plans to collect data following early discontinuations (of treatment and/or study)
    - But often without plans for how to use the post-discontinuation data collected
  - Increased wariness of single-valued imputation methods, (e.g., LOCF, BOCF)
    - Even as a sensitivity analysis, when such a strategy might be informative
  - Increased use of model-based methods, such as MMRM
    - But without sufficient attention to missing at random assumptions required for the analysis to be valid
OB Initiatives

• Provide sponsors with clarity about:
  – Our expectations for preventing missing data at the design stage and in study protocols
    • Including options for alternate study designs that may work in avoiding missing data but have limitations of their own
  – A better discussion about the treatment effect to be estimated in the regulatory context (the estimand), separate from the discussion about estimation methods for that effect
  – Regulatory expectations for the type and quantity of sensitivity analyses needed for regulatory decision-making
OB Initiatives

• External workshops (e.g., DIA/FDA Statistics Forum)
• Internal training at FDA
  – To ensure reviewers are on the same page with respect to missing data issues when engaging with sponsors and in reviews
  – To further understanding of clinical colleagues on missing data
• Recently submitted publications (3) that more broadly clarify FDA’s statistical policy for missing data
  – Prevention
  – Estimands
  – Sensitivity analyses
PREVENTION
Prevention

- Prevention of missing data is a key recommendation of the NRC report
- Application of the ITT principle calls for including as many randomized patients as possible in analysis
- Exclusion of some patients randomized to treatment, for reasons occurring after treatment began, may induce bias
- Missing data can be a "sign of a poorly conducted study"
Prevention

• What are the options for preventing missing data?
  – Consider alternative study designs
  – Create the right culture for patients
  – Avoid encouraging attrition
  – Plan for early discontinuations
Alternative Study Designs

- NRC comments on the need to consider alternatives to the parallel group design to avoid missing data
- Crossover studies can improve attrition with the promise of receiving experimental therapy for one of the periods
- Randomized-withdrawal studies are appealing for avoiding missing data, but
  - Not appropriate for all disease areas
  - Patient population differs (only responders are randomized)
  - Study may be under-powered, if effect size used for design is based on a traditional parallel group design
- Including a follow-on phase in which all patients receive experimental therapy can improve attrition, but
  - Efficacy results difficult to interpret from this phase, though tempting to try to do so
Prevention

• Create the right culture

  – Patients may feel they need to drop-out of the trial because their expectations of benefit are not met
  – Where possible, create a culture of patient care for the duration of the trial
  – Informed consent should accurately and clearly explain how patients will be managed if condition worsens, side effects occur, etc.
  – Informed consent can also invite patients to commit to furthering research, even if treatment is discontinued
Prevention

• Avoid encouraging attrition

  – Too many protocols still contain language instructing clinical centers to discontinue patients in violation of the protocol
    • For non-compliance with study medications, missed visits, etc.
    • Example protocol text: “If non-compliance during the treatment period continues, the subject may be terminated from the study at the investigator’s discretion.”
  – If patient’s safety is assured, non-compliance does not require early discontinuation
Prevention

• Plan for early discontinuations

  – Often know quite a lot about a particular disease and what to expect from the patients
  – Also may be able to anticipate tolerability issues from earlier phases of development
  – Take this knowledge into account when developing the study protocol and defining study endpoints
  – Emphasize importance of completeness and quality of data and standardize procedures for handling drop-outs as much as possible when training investigators and clinic personnel
Prevention

• Post-discontinuation data
  – It is usually the case that information on an early discontinuation case is useful
  – It is not always the case that post-discontinuation measurements can be used as the outcome for that case

• Are we doing the best job we can in ascertaining reasons for discontinuation?
  – Need objective ascertainment of reasons for early discontinuation
  – Case report form (CRF) design important
  – Training of clinic personnel is important
Prevention

• Example: Cystic Fibrosis (CF)
  – Chronic orphan disease afflicting ~30,000 US patients
  – Majority of patients treated by specialists in a network of CF clinics nationwide
    • Quarterly clinic visits for routine care
  – CF Foundation developed infrastructure to facilitate clinical trials conducted across the network (CF Therapeutic Development Network or CFTDN)
  – CFF also maintains patient registry with good representation of the patient population
CF as Example

• Example: Cystic Fibrosis (CF), cont.
  – Often the case that CF patients comprise a very compliant and engaged study population
  – CFTDN quality measures help minimize attrition for in-network therapeutic trials
  – Disease registry provides a back-up source for missing follow-up data
• Recent drug approval (ivacaftor) based on phase 3 studies with minimal attrition (<7%)
• When attrition is evident, relationship to treatment should be examined (see mannitol example)
Prevention

• Example at the other end of the spectrum: Schizophrenia
  – Extremely mobile population
  – May be anxious about discontinuing stable medication for entry into study
  – Early discontinuations often due to perceived lack of efficacy

• Therapeutic trials in this disease area with 30% attrition or higher not uncommon

⇒ Plan accordingly, where possible
Schizophrenia as Example

- **CATIE**: NIH-sponsored trial of atypical antipsychotics
  - Primary analysis = time to all-cause discontinuation
  - Other utility endpoints also considered combining efficacy and tolerability
- **CAMP**: NIH-sponsored trial, follow-on to CATIE
  - PANSS primary outcome
  - Post-discontinuation data collected but not used in primary analysis
  - Time to all-cause discontinuation key secondary endpoint

→ Interpretation of primary results on PANSS difficult in context of differential discontinuation
ESTIMANDS AND SENSITIVITY ANALYSIS
Estimands

• NRC report states that the trial protocol should define ‘the measures of intervention effects...’

• From a regulatory standpoint, it is not always clear in study protocols or analysis plans what is being estimated
  – More attention is typically given to the estimation methods

• With missing data, a discussion of the estimand becomes increasingly important

• Consider two hypothetical studies...
Hypothetical study #1

- Randomized, well-controlled clinical trial of a serious infectious disease
  - A single treatment is given by injection
  - Outcome is mortality at 28 days

- Some patients are lost to follow-up
  - And they are likely different from completers

- Prevention is the best solution – determine vital status on all patients, regardless of their status in the trial (*provided consent to do so is obtained*)

- Estimand is unambiguous – difference in proportions surviving
Hypothetical study #2

- Randomized, well-controlled clinical trial in diabetes
  - Treatment administered over six months
  - Outcome is HbA1c at end of study (Month 6) visit

- Some patients die during the study and do not have an HbA1c measurement at Month 6
  - Outcome is not missing; it does not exist

- Prevention for these patients is irrelevant

- Estimand is a composite endpoint of HbA1c and mortality
Hypothetical Studies

**#1 Survival unobserved**
1. Ascertainment failure
2. Missing outcomes exist
3. If ascertained, meaningful analysis would be straightforward
4. Numerous ascertainment failures may suggest poor study conduct
5. Imputations may be right or wrong
6. Missing value codes on dataset

**#2 HbA1c nonexistent**
1. Perfect ascertainment
2. Missing outcome does not exist
3. No meaningful, straightforward analysis
4. No failures, not poor conduct
5. Imputations are neither right or wrong
6. Missing value codes on dataset
Hypothetical study #3

- Randomized, well-controlled clinical trial in cystic fibrosis
  - Treatment administered over six months
  - Clinical outcome is long-term disease progression, morbidity, and mortality
  - Trial outcome is a surrogate endpoint, pulmonary function at Month 6

- The treatment has an unpleasant side effect (cough) which is also a sign of pharmacodynamics
  - And is the cause of patients dropping out of the trial

- Is this study more like #1 or #2?
Hypothetical study #3

- Like study #1:
  - Outcome exists--pulmonary function can be measured on drop-outs, even if they discontinue treatment
  - Estimand is difference in pulm fcn at Month 6 using retrieved values for drop-outs
  - Estimation methods needed (e.g., MMRM), if ascertainment is not fully successful
  - Analysis answers an important question – are patients treated with test drug better or worse off than those treated with control, regardless of whether they tolerated treatment for six months?
Hypothetical study #3

• But also like study #2:
  – Outcome may exist but may not be relevant
    • As a surrogate for long-term health effects, not being able to tolerate treatment makes pulmonary function at Month 6 irrelevant
  – The key outcome for drop-outs is the fact that they could not tolerate the treatment
  – Estimand is a composite endpoint combining pulmonary function and completion status (tolerability)
  – Estimation methods that impute values for drop-outs as if they could tolerate treatment are not meaningful here, just as imputing values that do not exist in study #2 would not be meaningful
Estimand summary

• If outcome measured after discontinuation is meaningful, then it should be measured or ascertained
  – Imputing or estimating it is not a substitute
  – If try and not possible, mathematically sophisticated methods are available for analysis

• If outcome after discontinuation is not meaningful, then it will also not be meaningful to estimate it
  – Because discontinuation is the outcome (as part of a composite)

• Another solution – estimate efficacy in an adherent subset
  – But need to synthesize a comparable subset of the control group (may be difficult)
Sensitivity Analysis

- Sensitivity analyses should be planned to assess the impact of missing data on the study results
  - Merely running additional analyses that make the same missing data assumptions is not useful
- For example, if a primary analysis assumes missing at random (e.g., MMRM), then a sensitivity analysis involving multiple imputation under the same assumption is uninformative
- The number of sensitivity analyses conducted is not as important as the way in which the assumptions are varied
- ‘Tipping point’ analysis recommended by NRC appealing for regulators
Example

January 2013 Advisory Committee* meeting to review mannitol inhalation powder for cystic fibrosis

- Highly significant results in one study; trend only in second study
- Successful study had differential drop-out rate by treatment, i.e., some patients could not tolerate therapy; second study did not
- Primary (pre-specified) analysis (test of average difference across visits in context of MMRM) not ideal—can’t define ‘tolerators’ in the control group, so comparison is flawed
- Sensitivity analysis included BOCF, but AC members skeptical of its utility
- FDA presented alternative sensitivity analyses to address tolerability (eCDF); suggested other analyses that may be needed

*http://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/Pulmonary-AllergyDrugsAdvisoryCommittee/ucm338461.htm.
Example

Differential Early Study Discontinuation

<table>
<thead>
<tr>
<th>Population</th>
<th>Study 301 (N=295)</th>
<th>Study 302 (N=305)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DPM</td>
<td>Control</td>
</tr>
<tr>
<td>ITT</td>
<td>176 (100%)</td>
<td>118 (100%)</td>
</tr>
<tr>
<td>Early study discontinuation (no post-baseline measurement)</td>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>MITT</td>
<td>156 (89%)</td>
<td>112 (95%)</td>
</tr>
<tr>
<td>Additional early study discontinuations</td>
<td>44</td>
<td>26</td>
</tr>
<tr>
<td>Completed 26 week treatment period</td>
<td>112 (64%)</td>
<td>86 (73%)</td>
</tr>
<tr>
<td>Reason for early study discontinuation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Withdrew by patient</td>
<td>28 (16%)</td>
<td>22 (19%)</td>
</tr>
<tr>
<td>Adverse event</td>
<td>29 (16%)</td>
<td>10 (8%)</td>
</tr>
<tr>
<td>Physician decision</td>
<td>6 (3%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Applicant decision</td>
<td>1 (&lt;1%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Other reasons</td>
<td>1 (&lt;1%)</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>
**Primary Efficacy Analysis using MMRM**

Change from Baseline in FEV$_1$ (mL) (MITT)

<table>
<thead>
<tr>
<th></th>
<th>DPM</th>
<th>Control</th>
<th>Treatment Comparison DPM - Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LS Mean (SE)</td>
<td>95% CI</td>
<td>p-value</td>
</tr>
<tr>
<td><strong>Average effect from week 6 to week 26 (LS mean (SE))</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study 301</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(m=156, c=112)*</td>
<td>118 (15)</td>
<td>35 (17)</td>
<td>83 (22)</td>
</tr>
<tr>
<td><strong>Study 302</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(m=177, c=120)*</td>
<td>107 (22)</td>
<td>52 (26)</td>
<td>54 (29)</td>
</tr>
</tbody>
</table>

*MITT population excludes subjects with no post-baseline data and statistical analysis methods assume missing data is missing at random*
**Example**

Sensitivity Analysis for Primary Endpoint (Baseline Observation Carried Forward)

Change from Baseline in $\text{FEV}_1$ (mL) (ITT)

<table>
<thead>
<tr>
<th></th>
<th>DPM</th>
<th>Control</th>
<th>Treatment Comparison</th>
<th>DPM - Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>BO CF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study 301</td>
<td>(m=176, c=118)</td>
<td>81 (14)</td>
<td>19 (18)</td>
<td>62 (24)</td>
</tr>
<tr>
<td>Study 302</td>
<td>(m=184, c=121)</td>
<td>76 (22)</td>
<td>12 (28)</td>
<td>65 (35)</td>
</tr>
</tbody>
</table>
Example

Post-hoc Analysis of Primary Efficacy Endpoint

Responder Analysis: Change from Baseline in FEV₁ (mL) at Week 26 (ITT)
Example #2

September 2013 Advisory Committee* meeting to review umeclidinium and vilanterol inhalation powder for COPD

- Differential drop-out rates in phase 3 studies – higher on placebo
- Statistical reviewer presented a sensitivity analysis using ‘jump to reference’ imputation assuming all drop-outs behaved as control group completers
  - Consistent with mechanism of action of drug
- Protocol did not call for post-treatment-discontinuation data collection
- Advisory Committee members questioned FDA’s criteria for missing data, i.e., how much is too much?
- FDA’s response: the reasons why data are missing is as important as the amount of data missing
- Impact of missing data on safety events (i.e., concern about under-reporting); no obvious imputation method to address this question
Summary: Analysis Planning

Key points for analysis planning:

1. The study design should be consistent with minimizing attrition, and study conduct should include efforts to retrieve meaningful data from drop-out cases

2. The main (primary) analysis has clearly stated assumptions and is valid under those assumptions
   - Pre-specification is not enough—assumptions should be plausible given the study population, study objectives, and anticipated effects of the test drug

3. Sensitivity analyses should be planned that correspond to alternative assumptions about missing data
References


EMA (2011), Guideline on Missing Data in Confirmatory Clinical Trials.


Some ITT classics:
- Friedman, Furberg, and DeMets, Fundamentals in Clinical Trials, 1985
References


References

Schizophrenia trials (NIH sponsored):


Recent CF drug approval: