Missing Data Handling in Non-Inferiority & Equivalence Trials
A systematic review

Brooke A Rabe
Graduate Interdisciplinary Program in Statistics
The University of Arizona, Tucson, USA

PSI’s Pharmaceutical Statistics Journal Club
Thursday, January 24, 2019
Acknowledgments

Simon Day  Clinical Trials Consulting & Training Limited, UK
Mallorie Fiero  US Food and Drug Administration
                Silver Spring, MD, USA.
Melanie Bell  Dept of Epidemiology and Biostatistics
                The University of Arizona, Tucson, USA.
Non-inferiority trials test the hypothesis that new treatment is not unacceptably worse than existing treatment (active control) by margin, $\delta$, (pre-specified).

- Used when experimental treatment not expected to be therapeutically superior to standard of care
- Has external benefits eg. cost, convenience of delivery, safety, or toxicity
- Cannot test directly against placebo

An increasing number of NI trials are being carried out
Suppose outcomes are continuous $\mu_T \equiv$ mean treatment effect in treatment group, $T$, $\mu_C \equiv$ mean treatment effect in control group, $C$, assuming larger values are better, the NI hypotheses take the form

\[ H_0 : \mu_C - \mu_T \geq \delta \] (T is inferior to C by $\delta$ or more)

\[ H_a : \mu_C - \mu_T < \delta \] (T is inferior to C by less than $\delta$, possibly superior)

(As though null and alternative hypotheses are switched.)
Conclude Non-Inferiority if Upper CL < Margin

(1) NI is demonstrated.
(2) NI is not demonstrated.
(3) NI is not demonstrated.
(4) NI is demonstrated but superiority to AC is not.
(5) NI and superiority demonstrated.
(6) NI is demonstrated but C is statistically superior to T

95% Confidence Intervals for C – T with NI margin, \( M_1 = 2 \).

\(^1\)US FDA, *Non-inferiority clinical trials to establish effectiveness; guidance for industry*, 2016.
Little guidance on missing data in NI trials
- NI extension of the CONSORT Statement does not discuss missing data handling in non-inferiority and equivalence trials
- National Research Council’s Report on missing data in clinical trials also does not discuss these types of trials

Missing data in superiority trials generally bias towards the null (conservative)

However, missing data in NI trials are generally anti-conservative
- Increases the likelihood of a type I error → claiming an inferior treatment is non-inferior
Analysis Sets in Non-Inferiority Trials

Selection of an appropriate analysis, intention-to-treat (ITT) or per-protocol, can be challenging.

- ITT analyses are generally preferred – they preserve randomization.
  - May be anti-conservative – missing data/non-compliance can weaken sensitivity to differences between groups – inflating type I error
- Are per-protocol analyses the answer? (Probably not.)
- Some suggest reporting both ITT and per-protocol
Non-inferiority trials lack internal validity (in contrast to superiority or placebo-controlled trials).

Are conditions in the non-inferiority trial sufficiently close to historic trials so that effect is preserved?

Need to ensure non-inferior conclusion implies superiority to placebo.

Missing data can weaken these assumptions.
Objectives of Systematic Review

1. Evaluate the extent, handling, sensitivity analysis for missing data and use of per-protocol, intention-to-treat analysis sets
2. Evaluate quality of reporting with respect to missing data
3. Collect data on sample size calculation, reasons why missing, relate quality of reporting to journal impact factor

Articles reviewed primarily by two reviewers, with six articles reviewed by both reviewers to establish consensus.

Oversampled from Annals of Internal Medicine, BMJ, JAMA, The Lancet, and The New England Journal of Medicine to compare with other systematic reviews of top journals.

After article selection, defined “high impact journals” as impact factor ≥ 5.
Main Results: Trial Characteristics

109 non-inferiority and equivalence trials reviewed

- 63% drug trials
- 38% justified the non-inferiority margin
- 54% accounted for missing data in sample size calculation
- 90% reported reasons why data were missing
- 35% continuous outcome
- 64% binary
- 1% count
Main Results: How Much Missing Data?

93% reported missing data in primary outcome.

Distribution of Proportion of Missing Data

- Proportion Missing Outcome at Individual Level
- Density

Brooke A Rabe
Missing Data in Non-Inferiority Trials
Main Results: Missing Data Handling

Stratified By Amount of Missing Data (< 10%, ≥ 10%)

<table>
<thead>
<tr>
<th>Method of handling missing data*</th>
<th>&lt; 10% (N = 60)</th>
<th>≥ 10% (N = 49)</th>
<th>Total (N = 109)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete case</td>
<td>32 (62%)</td>
<td>19 (39%)</td>
<td>51 (50%)</td>
</tr>
<tr>
<td>Single imputation**</td>
<td>14 (27%)</td>
<td>14 (29%)</td>
<td>28 (28%)</td>
</tr>
<tr>
<td>Unweighted GEE</td>
<td>1 (2%)</td>
<td>0 (0%)</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Multiple imputation</td>
<td>0 (0%)</td>
<td>6 (12%)</td>
<td>6 (6%)</td>
</tr>
<tr>
<td>Mixed model</td>
<td>1 (2%)</td>
<td>5 (10%)</td>
<td>6 (6%)</td>
</tr>
<tr>
<td>Unclear</td>
<td>4 (8%)</td>
<td>5 (10%)</td>
<td>9 (9%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analysis sets</th>
<th>&lt; 10% (N = 60)</th>
<th>≥ 10% (N = 49)</th>
<th>Total (N = 109)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITT or modified ITT only</td>
<td>24 (40%)</td>
<td>18 (38%)</td>
<td>42 (39%)</td>
</tr>
<tr>
<td>PP only</td>
<td>19 (32%)</td>
<td>13 (27%)</td>
<td>32 (29%)</td>
</tr>
<tr>
<td>Both</td>
<td>17 (28%)</td>
<td>18 (38%)</td>
<td>35 (32%)</td>
</tr>
<tr>
<td>Sensitivity analysis reported*</td>
<td>5 (10%)</td>
<td>6 (12%)</td>
<td>11 (11%)</td>
</tr>
</tbody>
</table>

*Of 101 articles with missing data reported. For sensitivity analyses, one trial explicitly made MNAR assumption but method was not specified. 7 used worst case imputation. **Last or baseline observation carried forward, best case, worst case.
### Main Results: By Trial Conclusion

<table>
<thead>
<tr>
<th></th>
<th>Conclusion of Non-inferiority</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes (N=90)</td>
</tr>
<tr>
<td>Sensitivity analysis reported</td>
<td>8 (10%)</td>
</tr>
<tr>
<td>Analysis sets</td>
<td></td>
</tr>
<tr>
<td>ITT or modified ITT only</td>
<td>35 (39%)</td>
</tr>
<tr>
<td>PP only</td>
<td>28 (31%)</td>
</tr>
<tr>
<td>Both</td>
<td>27 (30%)</td>
</tr>
<tr>
<td>Percent missing data</td>
<td></td>
</tr>
<tr>
<td>0-10%</td>
<td>51 (57%)</td>
</tr>
<tr>
<td>10-20%</td>
<td>29 (32%)</td>
</tr>
<tr>
<td>20-30%</td>
<td>5 (6%)</td>
</tr>
<tr>
<td>over 30%</td>
<td>5 (6%)</td>
</tr>
</tbody>
</table>
Secondary Results: Impact Factors

Most articles reviewed were from journals with IF < 5
- 31% from journals with IF > 10
- 16% from journals with $5 \leq \text{IF} \leq 10$
- 53% from journals with IF < 5

Analysis of both ITT and PP populations possibly more likely in high quality journals (IF > 5)
36% in high IF journals vs. 26% in other journals

Sensitivity analyses more likely in high quality journals
19% in high IF vs. 4% in other journals
### Secondary Results: Comparison to Other Reviews

<table>
<thead>
<tr>
<th></th>
<th>Individual RCT (Bell, 2013)</th>
<th>Cluster trials (Fiero, 2016)</th>
<th>Current study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trials with missing data</td>
<td>95%</td>
<td>93%</td>
<td>93%</td>
</tr>
<tr>
<td>Trials with &gt;10% missing data</td>
<td>47%</td>
<td>67%</td>
<td>45%</td>
</tr>
<tr>
<td>Complete case analysis</td>
<td>45%</td>
<td>55%</td>
<td>50%</td>
</tr>
<tr>
<td>Single imputation</td>
<td>27%</td>
<td>7%</td>
<td>28%</td>
</tr>
<tr>
<td>Multiple imputation/model-based</td>
<td>27%</td>
<td>29%</td>
<td>12%</td>
</tr>
<tr>
<td>Sensitivity analysis</td>
<td>37%</td>
<td>16%</td>
<td>11%</td>
</tr>
</tbody>
</table>
Conclusion

- Missing data in NI trials are common
- Complete case analysis and single imputation are popular
- Contrary to guidelines, sensitivity analyses are not being conducted (or reported)
- Minority of studies investigate both the ITT and PP populations
Recommendations

1. Analyze data with a likelihood-based approach or use multiple imputation under MAR assumption
2. May impute missing data under the null hypothesis
3. Conduct sensitivity analysis
4. Utilize repeated measures
5. Consider the primary estimand in making assumptions about imputed outcomes
Summary

- Missing data remain a problem in non-inferiority trials. In practice, their handling should be improved for more robust and reproducible clinical research.

- Missing data in non-inferiority trials need special consideration due to the anti-conservative nature of the design.

Ongoing and Future Work

- A conservative approach for intention-to-treat analysis of non-inferiority trials.

- Sensitivity analyses for missing data assumptions: proposing a tipping points approach using a pattern-mixture model in a multiple imputation framework.


