Disclaimer

The views expressed by the speaker during this seminar do not necessarily reflect the official views of the National Institute on Drug Abuse.
Slides are available on CTN’s Dissemination Library

Go to:

http://ctndisseminationlibrary.org/

and search for “statistical reasoning” or “wakim”
Any questions before we start?
Analysis Plan (Part 2 of 2)
Main Components of the Analysis Plan

1) ITT vs. per-protocol analysis
2) Statistical test or model
3) Multiplicity adjustment
4) Handling of missing data
5) Handling of outliers
6) Interim analyses
7) Sensitivity analysis
8) Secondary and subgroup analyses
Main Components of the Analysis Plan

1) ITT vs. per-protocol analysis
2) Statistical test or model
3) Multiplicity adjustment
4) Handling of Missing data
5) Handling of outliers
6) Interim analyses
7) Sensitivity analysis
8) Secondary and subgroup analyses
Multiplicity Adjustment

What is it?
When to do it?
Why do it?
What are the options?
When not to do it?
Multiplicity Adjustment: What is it?

Multiplicity adjustment is a way to control for false positive conclusions, i.e. to control the "family-wise" or "study-wise" rate of false positive conclusions.
Formally, whenever there are more than one primary endpoint (or primary hypothesis), more than two treatment conditions, more than one dose vs. placebo, or more than one time point

Informally, whenever there are more than one secondary analysis, including subgroup analyses
If you calculate many P values, some are likely to be small just by random chance. Therefore, it is impossible to interpret small P values without knowing how many comparisons were made. ... It is easy to be fooled by these small P values.

Mutolsky 2010
Recall:

Alpha = Type I error = chance of finding a statistically significant result when the null hypothesis is true (e.g. no difference)

Alpha = 0.05 is most commonly used
Multipliclity Adjustment: Why do it?

Example:

Two endpoints: drug use and retention

Each endpoint is tested at the 5% alpha level

The experimental treatment is considered beneficial if either or both endpoints are significant

Without multiplicity adjustment, the chance of the treatment being found beneficial when it is not can be as high as 10% (not 5%)

Clinical Trials Network
National Institute on Drug Abuse — National Institutes of Health — U.S. Department of Health and Human Services
Multiplicity Adjustment: What are the options?

1) Basic Procedures

2) Stepwise Procedure (*pre-specified* testing sequence)

3) Stepwise Procedure (*data-driven* testing sequence)

4) Other more complicated methods

Dmitrienko 2011
1) Basic Procedures, e.g. Bonferroni:

P-values are compared to a pre-specified fraction of the alpha level (0.05).

Example: 3 tests $\rightarrow$ new alpha $= 0.05/3 = 0.017$

Pros: simple

Cons: least powerful (most conservative)
Multipliclity Adjustment: What are the options?

2) Stepwise Procedure (pre-specified testing sequence) e.g. fixed-sequence procedure.

Hypotheses are ordered a priori, typically reflecting clinical importance.

Testing begins with the first hypothesis, and each test is carried out without a multiplicity adjustment as long as significant results are observed in all preceding tests, i.e. the testing stops when the first non-significant result is observed.

Dmitrienko 2011 & Dmitrienko 2009
3) Stepwise Procedure (data-driven testing sequence), e.g. Holm, Hochberg & Hommel

Start with the lowest p-value (Holm) or highest p-value (Hochberg & Hommel) and follow a sequence of steps

Hommel’s is more powerful than Hochberg’s, which is more powerful than Holm’s

Dmitrienko 2011
Multiplicity Adjustment: When *not* to do it (i.e. not necessary)?

When *all* the primary endpoints have to be statistically significant in order to claim treatment benefit, e.g. to get FDA approval.

EMEA/CPMP 2002

Example: the experimental treatment is considered beneficial only if *both* drug use *and* retention are found to be statistically significant.
Main Components of the Analysis Plan

1) ITT vs. per-protocol analysis
2) Statistical test or model
3) Multiplicity adjustment
4) Handling of Missing data
5) Handling of outliers
6) Interim analyses
7) Sensitivity analysis
8) Secondary and subgroup analyses
The Prevention and Treatment of Missing Data in Clinical Trials
Panel on Handling Missing Data in Clinical Trials
National Research Council of the National Academies
July 2010

8 recommendations on minimizing missing data

12 recommendations on statistical approaches
Extent of the Issue (in the CTN)

Based on the first 24 multi-site clinical trials on substance abuse conducted between 2001 and 2010 in NIDA’s Clinical Trials Network (CTN), the percent of missing data for the primary outcome measure ranged from 2% to 60% (median=25%).

Wakim et al. 2011
What’s the big deal?

We need N=450 (based on power analysis)

And we expect 25% missing

So we set the initial N=600

So that the final (analyzed) N=450
Technical terms that we can’t escape...

- Missing at random (MAR)
- Missing completely at random (MCAR)
- Missing not at random (MNAR)

Ignorable

Non-ignorable

... but what do they mean?
(Non-technical) Definition:
The fact that Y is missing has nothing to do with its unobserved value, or with other measured variables.

Therefore:
The set of participants with complete data can be regarded as a simple random (or representative) sample of all participants.

What to do?
Ignore the missing data and analyze the available data ("complete case" or "pairwise deletion" method).
**Missing at Random (MAR)**

(Non-technical) Definition:
The fact that Y is missing *can* be explained by other values of Y, or by other measured variables.

Therefore:
The observed data can be used to account for the missing data.

What to do?
Use Maximum Likelihood or Multiple Imputation approach, and include in the model the other measured variables that explain missingness.
(Non-technical) Definition:
The fact that Y is missing *cannot* be explained by other values of Y, or by other measured variables

Therefore:
The observed data cannot be used to account for the missing data; and outside information is needed

In simple English:
We have a problem
## In Summary...

<table>
<thead>
<tr>
<th>Missingness (i.e. whether the data are missing or not)</th>
<th>is related to</th>
<th>is not related to</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCAR</td>
<td></td>
<td>observed or unobserved data</td>
</tr>
<tr>
<td>MAR</td>
<td>observed data</td>
<td>unobserved data</td>
</tr>
<tr>
<td>MNAR</td>
<td>unobserved data</td>
<td></td>
</tr>
</tbody>
</table>

Based on Graham 2009
**Bottom Line**

**MCAR:** No big deal

**MAR:** Use available collected data to “explain” missing mechanism, and use existing statistical methods

**MNAR:** Need outside information to “explain” missing mechanism
Ignorable & Non-Ignorable
(roughly speaking)

Ignorable (available data is sufficient):
• Missing Completely At Random (MCAR)
• Missing At Random (MAR)

Non-Ignorable (need outside information):
• Missing Not At Random (MNAR)
Main Components of the Analysis Plan

1) ITT vs. per-protocol analysis
2) Statistical test or model
3) Multiplicity adjustment
4) Handling of Missing data
5) Handling of outliers
6) Interim analyses
7) Sensitivity analysis
8) Secondary and subgroup analyses
Handling of Outliers

What is an outlier?
How do outliers arise?
How are outliers identified?
Is it legitimate to remove outliers?
How should outlier removal be reported?

Motulsky 2010
What is an outlier?

An outlier is a value that is so far from the others that it appears to have come from a different population.

The presence of outliers can invalidate many statistical analyses.

Motulsky 2010
How do outliers arise?

Incorrect value

- Invalid data entry
- Experimental mistakes
- Biological diversity
- Random chance
- Wrong assumption

Correct value

Motulsky 2010
How are outliers identified?

- Statistical tests
- Single vs. multiple outliers
- Ultimately a subjective exercise

Motulsky 2010
Is it legitimate to remove outliers?

When is it “cheating” and when is it the responsible thing to do?

It’s all about pre-specification and disclosure

Motulsky 2010
How should outlier removal be reported?

- Keep the outlying observations in the database, with a flag
- Show a graph with all values, and the outliers identified/marked
- Report how many outliers were excluded from the primary analysis, and the criteria used to identify the outliers
- Consider reporting the results in two ways: with and without the outliers

Motulsky 2010
Main Components of the Analysis Plan

1) ITT vs. per-protocol analysis
2) Statistical test or model
3) Multiplicity adjustment
4) Handling of Missing data
5) Handling of outliers
6) Interim analyses
7) Sensitivity analysis
8) Secondary and subgroup analyses
Main Components of the Analysis Plan

1) ITT vs. per-protocol analysis
2) Statistical test or model
3) Multiplicity adjustment
4) Handling of Missing data
5) Handling of outliers
6) Interim analyses
7) Sensitivity analysis
8) Secondary and subgroup analyses
References


FDA, Guidance document on multiplicity issues in clinical trials, expected to be released in December 2011 (?).


Questions or Comments