Statistical Considerations for Human Abuse Potential Studies

Ling Chen, Ph.D.
FDA/CDER/OB/DBVI

The 11th Annual Scientific Meeting, Washington, DC
February 18, 2015
Disclaimer

This presentation reflects the views of the author and should not be construed to represent FDA’s views or policies.

Note: The human abuse potential (HAP) study is also called the human abuse liability (HAL) study.
Outline

• The HAP study: an efficacy study or a safety study
• Multiplicity issue
• Missing data issue for sedative products
• Subgroup analysis
• Concluding remarks
Efficacy or Safety Study

• General HAP studies
  The assessment of abuse potential of the test drug in the general HAP study is to look for abuse potential signals, which is a part of the safety profile assessment of the test drug. Thus, such a study is a type of safety studies.

• HAP studies for reformulated opioid products
  The study is for assessing the abuse deterrent effects for the reformulated opioid product relative to the marketed opioid product. It can be considered as an efficacy study, because the study evaluates if the test drug produces a desired effect.

Note: This presentation focuses on the general HAP studies.
A General HAP Study

• The HAP study was a randomized, double-blind, placebo- and positive-controlled, 7-way crossover study to evaluate the abuse potential of single doses of the test drug compared to placebo, drug X, and drug Y, in healthy male and female recreational polydrug users.
• X is a stimulant drug, and Y is a sedative drug.
• Polydrug user refers to the use of two or more psychoactive drugs in combination to achieve a particular effect. Thus, it is expected that these subjects would respond to both the sedative and the stimulant.
Comparisons
When can the type I error rate inflate?

• A rule of thumb in efficacy studies:
  – If a clinical decision rule for efficacy poses multiple opportunities to win, then generally there can be Type I error rate inflation requiring adjustments for multiplicity.

• Examples:
  – Epilepsy trial: win on Seizure rate or
    (win on Drop Attack Rate and win on Seizure Severity)
  – Multiple statistical tests or comparisons between the test drug (for example, three doses) and placebo in an efficacy study.

The principle of adjustments for multiplicity in efficacy studies can be extended to HAP studies.
Adjustments for Multiplicity?

Study Validation

Claim for no abuse potential signal

Claim for less abuse potential than positive controls

Dose response

NO MULTIPLICITY ADJUSTMENT IS REQUIRED FOR ANY COMPARISON IN GENERAL HAP STUDIES!
Co-primary endpoints

**Definition** Two or more specified primary endpoints are co-primary, when each must show that there is a statistically significant beneficial effect of the experimental treatment.

**Example** (Alzheimer’s trials in mild to moderate disease)

1. ‘ADAS-Cognitive Subscale’ endpoint: measures the change in the patient’s cognitive function
2. Clinician’s Interview Based Impression of Change (CIBIC) endpoint: measures deficits of activities of daily living.

Both endpoints must show statistically significant benefit of the study treatment.

Primary Measures

Drug Liking VAS

High VAS

ARCI PCAG

Take Drug Again VAS
Co-primary endpoints

• The co-primary endpoints in this study are
  - Emax on Drug Liking VAS
  - Emax on High VAS
  - Emax on ARCI PCAG
  - Emax on Take Drug Again VAS

• Good News:
  No adjustment of the type I error rate for the single co-primary endpoint is required.

Bad News?
### % Increase on Sample Size

<table>
<thead>
<tr>
<th>Correlation</th>
<th>Number of co-primary endpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>0.2</td>
<td>29</td>
</tr>
<tr>
<td>0.5</td>
<td>25</td>
</tr>
<tr>
<td>0.8</td>
<td>17</td>
</tr>
</tbody>
</table>

- Assume same effect size on each endpoint, tested at one-sided 2.5% level. The objective is to have an 80% overall power.

Limit primary abuse potential measures of interest

• Before a HAP study, the sponsor should gather the information about abuse potential of the test drug, for example, from receptor binding studies, animal behavioral studies, and the evaluation of adverse events that occur during clinical studies, and then discuss with the CSS for choosing limited the primary abuse potential measures of interest in the study. The influential statistics should be based on co-primary endpoints. For endpoints on other measures, the descriptive statistics should be provided in the study report.
Reduce the Number of Comparisons

• Do the validation tests first.
• Do dose response tests second.
• Suppose the T2 has the largest mean liking in Emax, the mean liking in Emax of Y1 is smaller than that of Y2. Then

Claim for less abuse potential than positive controls
Claim for no abuse potential signal

Original 12 comparisons can be reduced to 3 comparisons!
Missing Data Issue with Sedative Products

Individual time course profiles for Drug Liking VAS (the high dose test drug)

• Grey on the heat map shows missing data: 5.9% (2/34), 32.4% (11/34), 26.5% (9/34), and 23.5% (8/34) of subjects had missing data at hours 1, 2, 3 and 4, respectively.

• The AE report shows that 15, 3, and 1 subjects experienced somnolence, drowsiness and confusion, respectively.

• Because the primary endpoint is Emax, the missing data issue was not noticed until the heat map was developed.

• How should we deal with missing data issue?
Continued…

• Imputation for missing data is not needed for this case.
  - The primary endpoint is based on Emax. Using existing data of a subject to impute the subject’s missing data would not change Emax.
  - One should not use average or other statistics at a time point from the subjects who did not have missing data to impute the missing data at this time point, because subjects who had missing data in this case is a subgroup of the study population which is different from other subjects in the study.

• Subjects who have missing data due to sedative effect should not be excluded from the statistical analysis.
  - These subjects are part of the study population.
  - However, a completer in the HAP study should have at least one observation around $t_{max}$ for Drug Liking VAS for each treatment in the study.
Subgroup Analysis

- In this example, females have much longer half-life than males (130/84 hours). The assessment on gender difference was one of the interests in this study, particularly, the abuse potential of the test drug for females.

- Because this is a sedative drug, studying the difference between the asleep group and the awake group may also be of interest, particularly, the abuse potential of the awake group.

- In a HAP study, if a subgroup is of interest, and a claim is to be made for a subgroup in addition to that of the overall study population, no adjustment for multiplicity needs to be made to control type I error rate, because the sponsor needs to win either “no abuse potential” or “less abuse potential” of the test drug for both overall study population and the subgroup.

- However, the study needs to be powered for the subgroup in addition to the overall population for the subgroup analysis.
Concluding Remarks

- Even though there are multiple comparisons and co-primary endpoints in general HAP studies, adjustments for multiplicity are not needed. However, the sample size calculation should take into account for multiple comparisons and co-primary endpoints to ensure a proper power of the study.
- Missing data due to sedative effect do not need to be imputed. However, a completer should have at least one observation around $t_{\text{max}}$ for Drug Liking VAS for each treatment in the study.
- If a subgroup analysis is needed in a general HAP study, an adjustment for multiplicity for the subgroup analysis is not needed. However, the study needs to be powered for the subgroup in addition to the overall population.