An Overview of Regulatory Recommendations for Statistical Approaches to Evaluate Human Abuse Potential Study Data

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

Advancement in Abuse Potential Assessments – Building on the FDA Draft Guidance for Industry, April 16-17, 2015

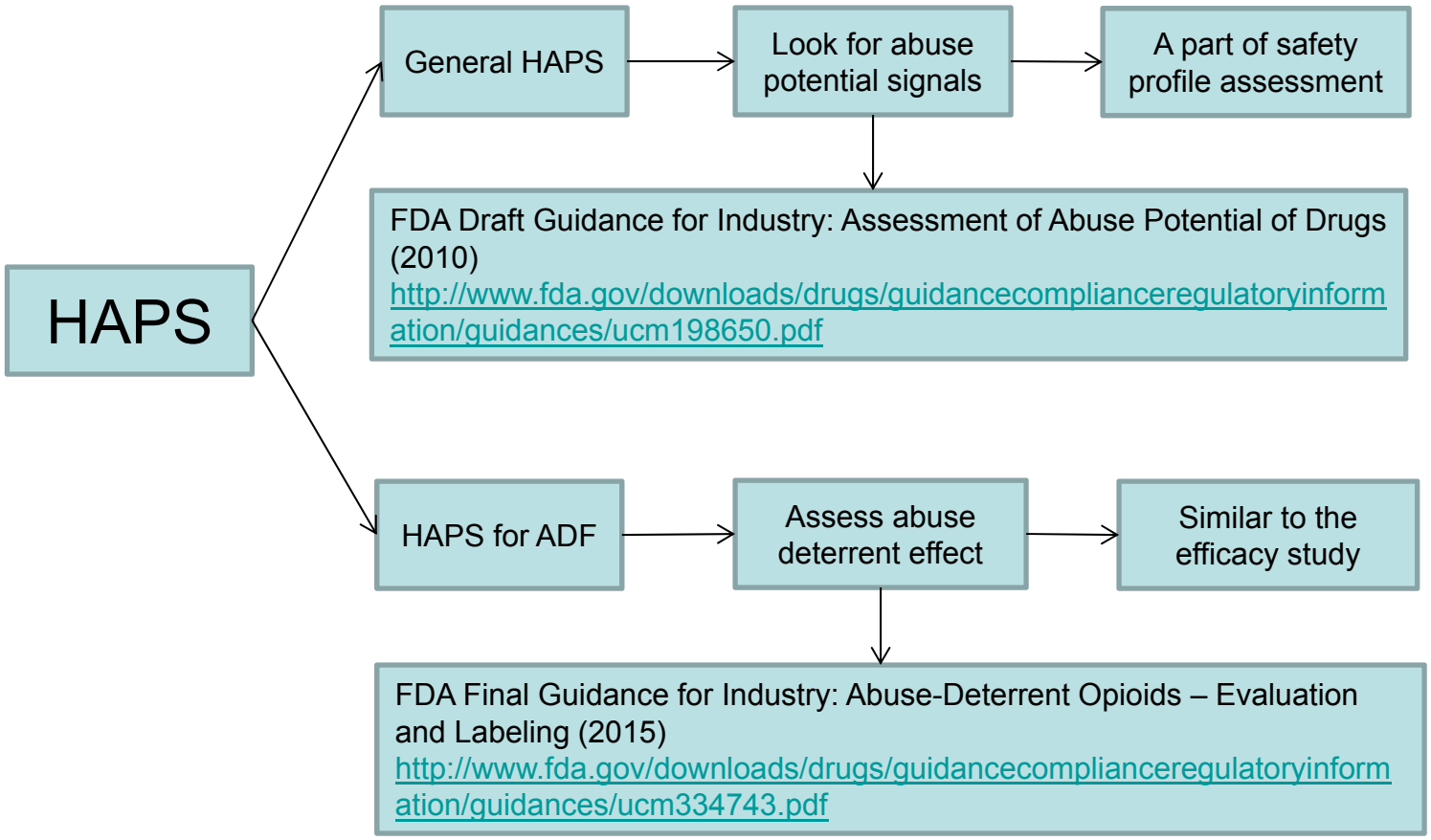
Disclaimer

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

Outline

- The human abuse potential study (HAPS): an efficacy study or a safety study?
- Hypotheses in HAPS
- Multiple comparisons and co-primary endpoints
- Statistical tests
- Summary

Efficacy or Safety Study



Hypotheses

1. The comparison between the positive control and placebo

   $H_0 : \mu_C \leq \mu_P$ versus $H_a : \mu_C > \mu_P$ (Study validation)

2. The comparison between the test drug and the positive control

   $H_0 : \mu_T \geq \mu_C$ versus $H_a : \mu_T < \mu_C$

3. The comparison between the test drug and placebo

   $H_0 : \mu_T - \mu_P \geq 11$ versus $H_a : \mu_T - \mu_P < 11$

Equivalence Margin

All tests are one-sided and at the 2.5% significance level.

When can the type I error 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.
- The principle of adjustments for multiplicity in efficacy studies can be extended to HAPS.

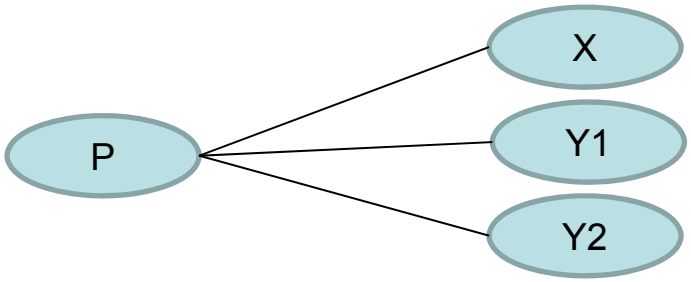
Example

- There are 7 treatments in the HAPS.

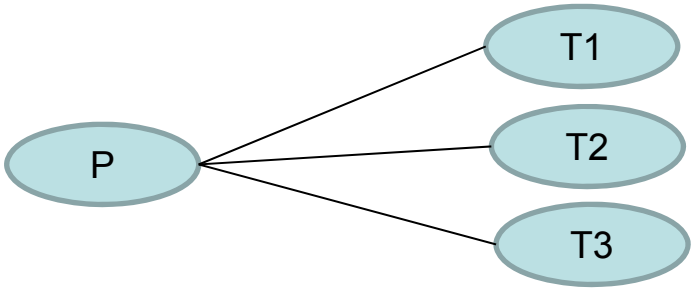
X – Positive control (stimulant)
Y1 – Low dose positive control (sedative)
Y2 – High dose positive control (sedative)
T1 – Low dose test drug
T2 – Medium dose test drug
T3 – High dose test drug
P – Placebo

Adjustments for Multiplicity?

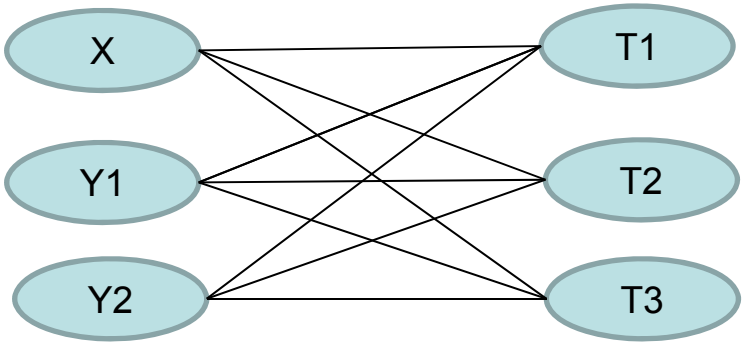
Study Validation



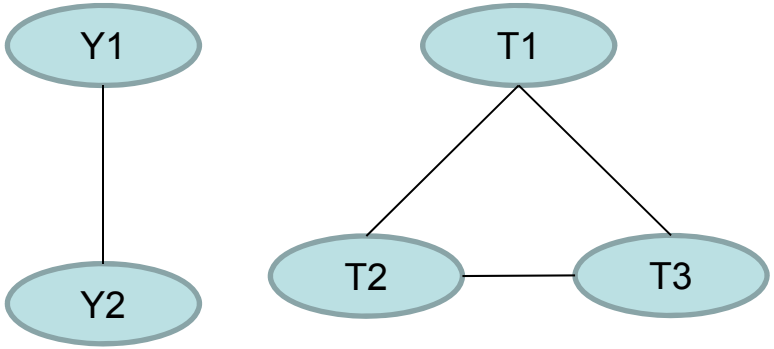
Claim for no abuse potential signal



Claim for less abuse potential than positive controls

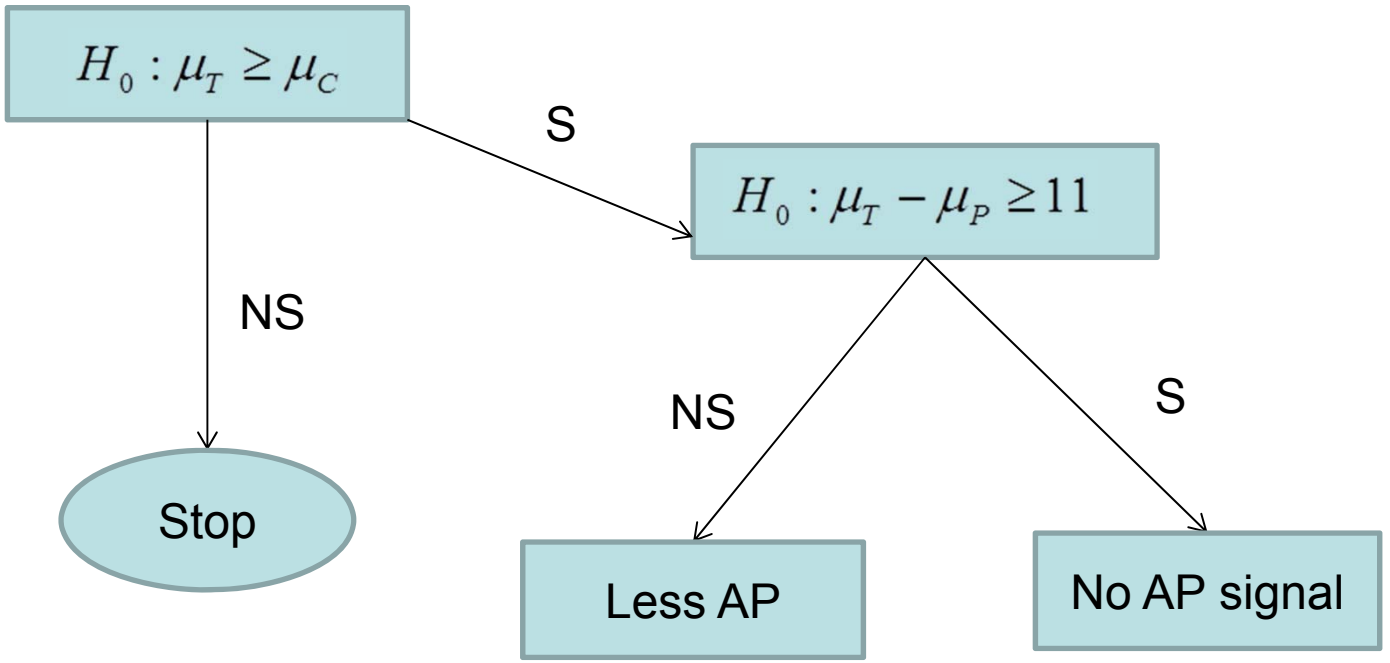


Dose response



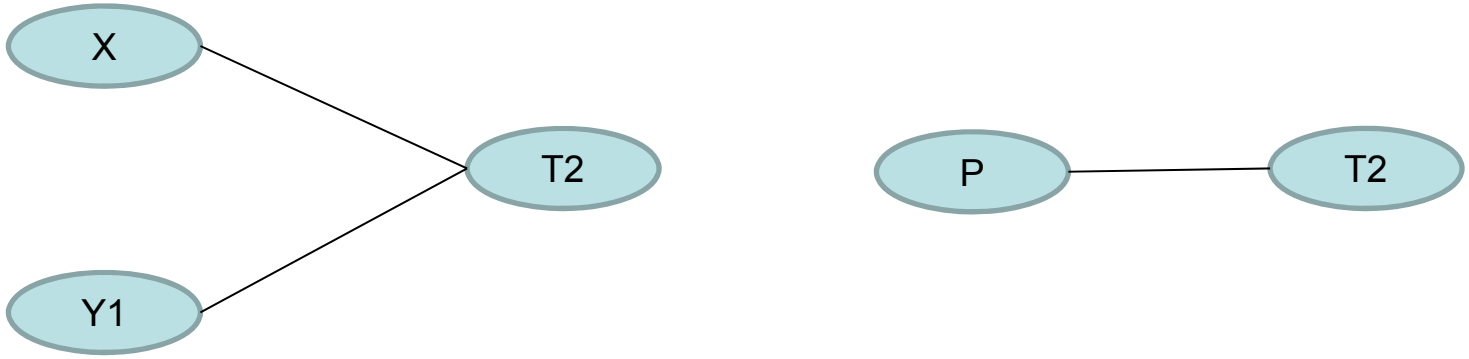
NO MULTIPLICITY ADJUSTMENT IS REQUIRED FOR ANY COMPARISON IN GENERAL HAPS!

Closed Testing Procedure



Reduce the Number of Comparisons in the Primary Analysis

- 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



Original 12 comparisons can be reduced to 3 comparisons!

Co-primary Endpoints

- Definition (efficacy studies) 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

- Primary measures: Drug Liking VAS, High VAS, Take Drug Again VAS
- Endpoints of interest: Emax, AUE_{0-2}, and T_{Emax}
- There are 9 co-primary endpoints: Drug Liking VAS (Emax, AUE_{0-2} and T_{Emax}), High VAS (Emax, AUE_{0-2} and T_{Emax}), and Take Drug Again VAS (Emax, AUE_{0-2} and T_{Emax}).
- Good news: No adjustment of the type I error rate for the single co-primary endpoint is required.
- Bad news?

% Increase on Sample Size

	Number of co-primary endpoints			
Correlation	2	3	4	9
0	**31**	**49**	**62**	**96**
0.2	**29**	**46**	**58**	**91**
0.5	**25**	**39**	**49**	**74**
0.8	**17**	**27**	**32**	**48**

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

Source: Christy Chuang-Stein, "Challenge of multiple co-primary endpoints: A new approach", the 2007 ICSA Applied Statistics Symposium.

The Wilk-Shapiro test (The W test)

- Before performing the comparisons, one should examine the normality assumption of the statistical model by using the W test.

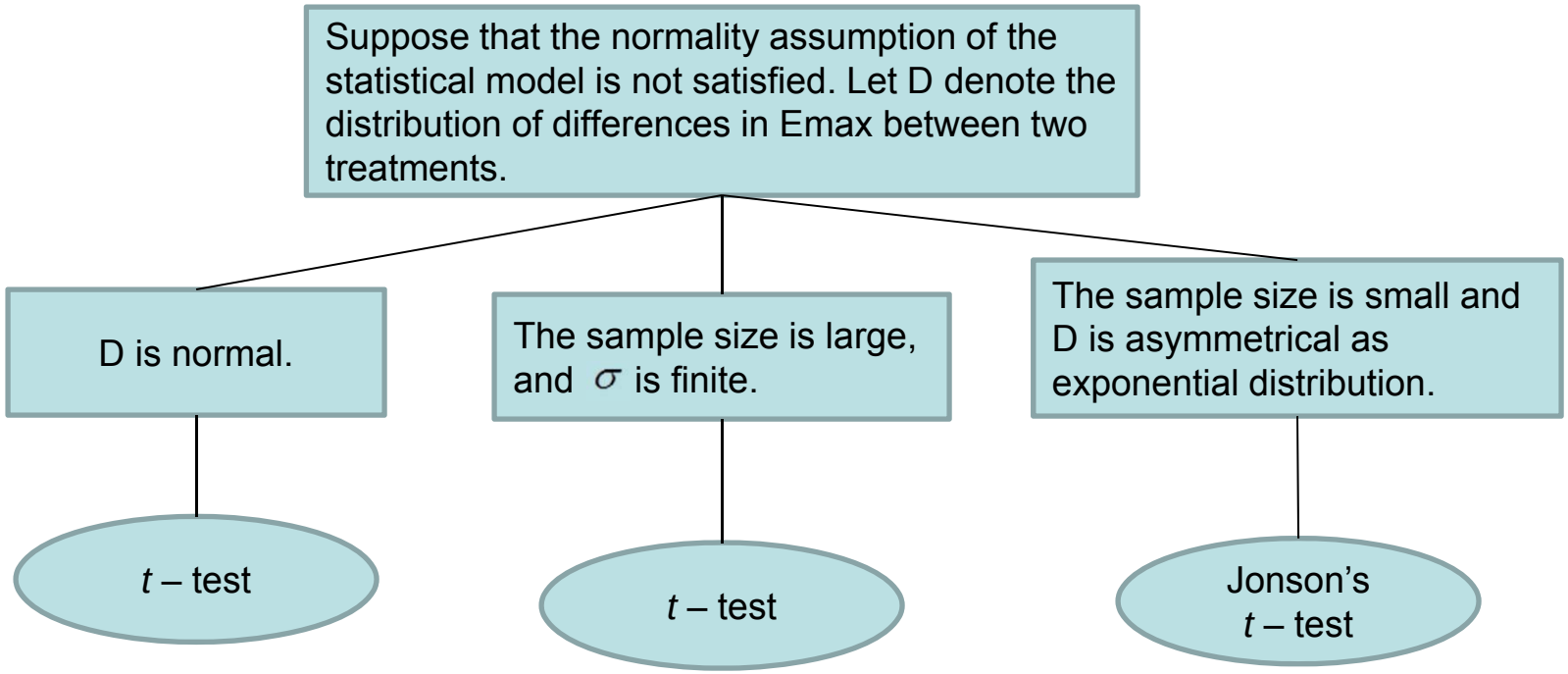
- Suppose the p-value of the W test for residuals is 0.0005. Should we conclude that the residuals are normally distributed?

  H_0: The distribution of residuals is normal

  H_a: The distribution of residuals is not normal

Therefore, based on p-value=0.0005, there is a strong evidence that the residuals are not normally distributed.

Statistical Tests for Comparisons



- If D is very skewed, you may use the normal approximation for the sign test to test the median difference (n≥12) or pre-specify an alternate method of this test in the SAP.
- Note that Wilcoxon-Signed Rank test has an assumption that the paired differences all come from the same continuous, symmetric distribution.

Summary

- Do not need multiplicity adjustments for type I error rate for multiple comparisons and co-primary endpoints in general HAPS.
- The sample size calculation should take into account for the multiple comparisons and co-primary endpoints.
- Pay attention to hypotheses and assumptions for proper statistical tests.

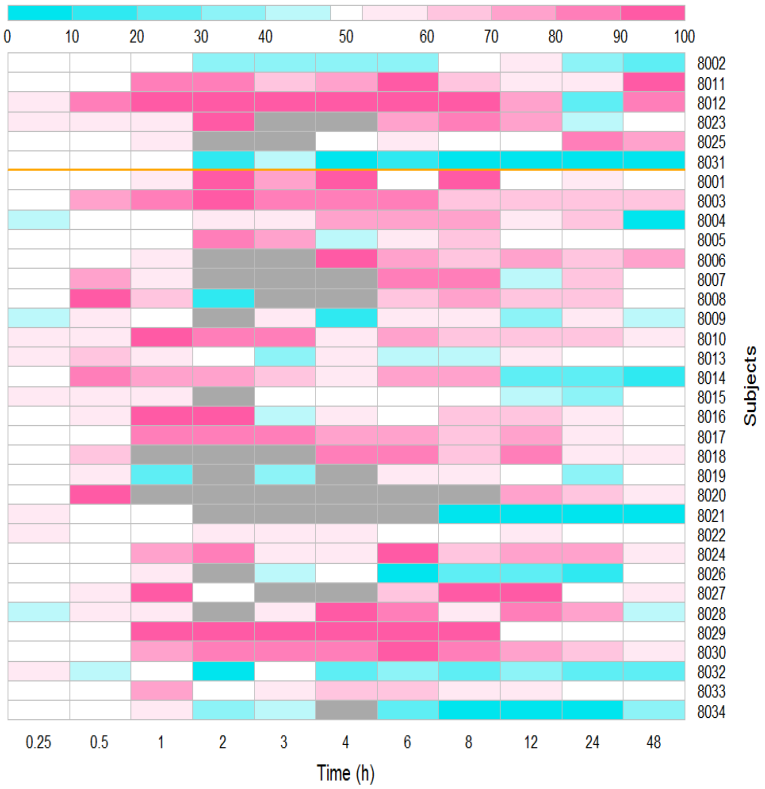
References

1. Chen, L. and Bonson, K. R. (2013). "An Equivalence Test for the Comparison between a Test Drug and Placebo in Human Abuse Potential Studies," Journal of Biopharmaceutical Statistics, 23: 294-306.
2. Marcus, R; Peritz, E; Gabriel, KR (1976). "On closed testing procedures with special reference to ordered analysis of variance". Biometrika 63: 655–660.
3. Shapiro, S. S. and Wilk, M. B. (1965). "An analysis of variance test for normality (complete samples)". Biometrika, 52 (3–4): 591–611
4. Chen, L. & Shapiro, S. S. (1995). "An Alternative Test for Normality Based on Normalized Spacings," Journal of Statistical Computation and Simulation, 53: 269-288.
5. "Student" (1908). "The probable Error of a Mean," Biometrika, 21, 259-286.
6. Miller, R. G., Jr. (1986). "Beyond ANOVA: Basic of Applied Statistics, New York: John Wiley.
7. Johnson, N.J. (1978). "Modified t Tests and Confidence Intervals for Asymmetrical Populations," Journal of the American Statistical Association, 73, 536-544.
8. Chen, L. (1995). "Testing the Mean of Skewed Distributions," Journal of the American Statistical Association, 90: 767-772.
9. Hollander, M. and Wolfe, D. A. (1973). Nonparametric Statistical Methods. New York: John Wiley & Sons.
10. Lehmann, E. L. (1975). Nonparametrics: Statistical Methods Based on Ranks. San Francisco: Holden-Day.

Thank you!

Missing Data Issue with Sedative Products

Individual time course profiles for Drug Liking VAS (A sedative product)



- Imputation for missing data is not needed for this case.
  - 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 from awaked subjects to impute missing data. This is against the principle of the crossover 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 HAPS should have at least one observation around t_{max} for Drug Liking VAS for each treatment in the study.