CCALC

Cross Company Abuse Liability Council

Statistical Issues in Abuse-Deterrent Formulation (ADF) and Human Abuse Potential (HAP) Studies

Catherine Mills

Principal Data Analyst, Early Phase, Syneos Health

Abuse Potential Dialogue Session

12-Oct-2018

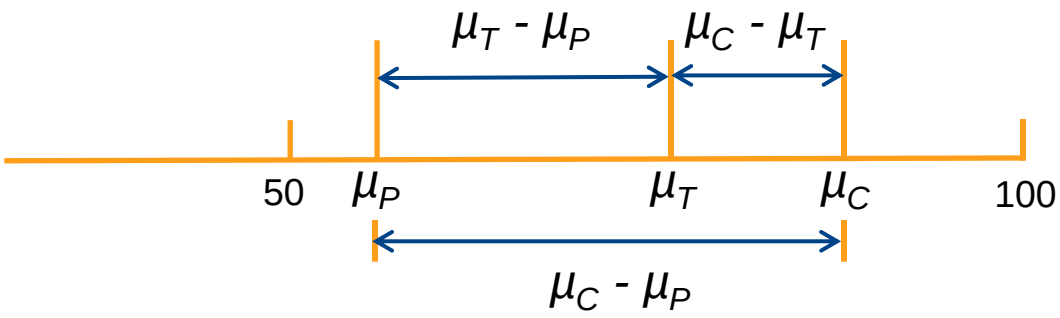
Disclosure

I am an employee of Syneos Health and
in my role I consult with various pharmaceutical and biotech companies

Overview

- Hypothesis Testing for Abuse-Deterrent Opioid Studies - Margin Selection
  - FDA Final Guidance, April, 2015
- Hypothesis Testing for Human Abuse Potential (HAP) Studies - Margin Selection
  - FDA Draft Guidance, January, 2010
  - FDA Final Guidance, January, 2017
- Sample Size Calculations
- Qualification Data vs. Treatment Data

Drug Liking VAS E_{max}: Treatment Differences of Interest



- P = Placebo; C = Positive Control; T = Test Drug

Pre-Final Guidance Approach

1. Does the positive control (C) produce mean responses that show different abuse potential compared to placebo (P)?

   $H_0: \mu_C - \mu_P = \delta_1$ *vs.* $H_a: \mu_C - \mu_P \neq \delta_1$, $\delta_1 = 0$

2. Does the test drug (T) produce mean responses that show different abuse potential compared to positive control?

   $H_0: \mu_C - \mu_T = \delta_2$ *vs.* $H_a: \mu_C - \mu_T \neq \delta_2$, $\delta_2 = 0$

3. Does the test drug produce mean responses that show different abuse potential compared to placebo?

   $H_0: \mu_T - \mu_P = \delta_3$ *vs.* $H_a: \mu_T - \mu_P \neq \delta_3$, $\delta_3 = 0$

Final Guidance Hypotheses
FDA Margin Recommendations for Drug Liking VAS E_{max}

1. Does the positive control (C) produce mean responses that show greater abuse potential compared to placebo (P)?

$$H_0: \mu_C - \mu_P \leq \delta_1 \textit{ vs. } H_a: \mu_C - \mu_P > \delta_1, \qquad \delta_1 > 0; \delta_1 = 15$$

2. Does the test drug (T) produce mean responses that show less abuse potential compared to positive control?

$$H_0: \mu_C - \mu_T \leq \delta_2 \textit{ vs. } H_a: \mu_C - \mu_T > \delta_2, \qquad \delta_2 \geq 0; \delta_2 = 0$$

3. Does the test drug produce mean responses that show similar abuse potential compared to placebo?

$$H_0: \mu_T - \mu_P \geq \delta_3 \textit{ vs. } H_a: \mu_T - \mu_P < \delta_3, \qquad \delta_3 > 0; \delta_3 = 11$$

Margin Selection for Hypothesis Testing

"The actual values of δ_1, δ_2, and δ_3 vary according to such factors as subjective measures, drug class, and route of drug administration.

All the margins should be pre-specified and justified in the protocol.

The statistical tests yield multiple comparisons (all doses of positive control drug versus placebo; all doses of the test drug versus each dose of the positive control drug; and all doses of the test drug versus placebo) for each of the subjective measures collected.

For each hypothesis, the statistical significance of the test should be achieved on all doses.

Thus, no multiplicity adjustment is recommended."

FDA Final Guidance, January 2017

Sample Size Calculations for Hypothesis 1: Positive Control vs. Placebo – Oral HAP Studies

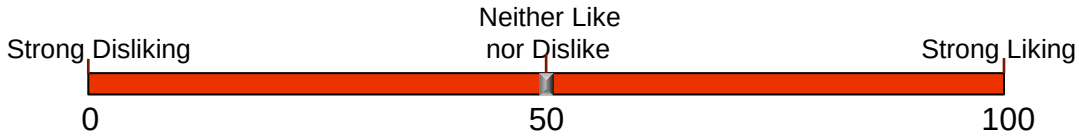
					$H_o: \mu_C - \mu_P = 0$ versus $H_a: \mu_C - \mu_P \neq 0$	$H_o: \mu_C - \mu_P \leq 15$ versus $H_a: \mu_C - \mu_P > 15$
Contrast (Differences)	Mean Difference	SD	Within-Subject Correlation	Power (%)	Sample Size from SAS® PROC POWER	Sample Size from SAS® PROC POWER
Phentermine 45 mg n=37, alpha=0.05	18.0	19.28	0.0291	90	26	**677**
Phentermine 60 mg n=38, alpha=0.05	22.7	17.65	0.0671	90	14	**87**
Zolpidem 30 mg n=32, alpha=0.05	20.5	21.15	0.0995	90	23	**227**
Suvorexant 40 mg n=32, alpha=0.05	18.3	24.41	-0.2408	90	49	**1156**
Alprazolam 1.5 mg n=39, alpha=0.05	27.2	13.1	0.2429	90	6	17
Alprazolam 3 mg n=39, alpha=0.05	33.0	12.66	0.4294	90	5	7
Lorazepam 2 mg n=34, alpha=0.05	19.4	15.03	0.3065	90	11	**138**
Lorazepam 4 mg n=34, alpha=0.05	23.9	14.59	0.1001	90	10	43
Marinol 10 mg n=33, alpha=0.05	23.9	23.80	-0.4034	90	32	**175**
Marinol 30 mg n=33, alpha=0.05	34.8	15.92	0.0355	90	7	13

Sample Size Calculations for Hypothesis 1: Positive Control vs. Placebo – Intranasal ADF Studies

					$H_o: \mu_C - \mu_P = 0$ versus $H_a: \mu_C - \mu_P \neq 0$	$H_o: \mu_C - \mu_P \leq 15$ versus $H_a: \mu_C - \mu_P > 15$
Contrast (Differences)	Mean Difference	SD	Within-Subject Correlation	Power (%)	Sample Size from SAS® PROC POWER	Sample Size from SAS® PROC POWER
Stimulant 40 mg n=37, alpha=0.025	22.2	20.49	0.0579	90	19	**163**
Oxycodone HCl 15 mg n=42, alpha=0.025	38.1	19.44	0.0670	90	8	16
Oxycodone HCl 30 mg n=37, alpha=0.025	40.6	11.68	-0.0210	90	5	7
Hydrocodone 50 mg n=21, alpha=0.025	33.7	18.04	0.2068	90	7	18

Qualification Criteria

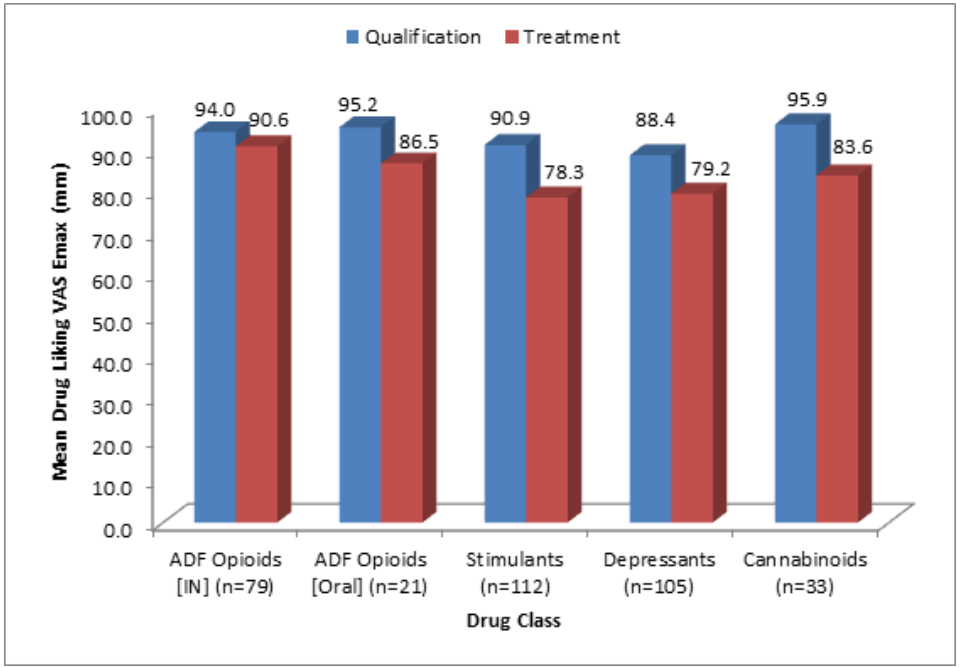
At this moment, my liking for this drug is



1. Able to distinguish active comparator(s) from placebo on a bipolar Drug Liking (At this Moment) VAS, defined as ≥15 point peak increase for Drug Liking in response to each active comparator relative to placebo
2. Peak score of ≥ 65 on Drug Liking VAS in response to the active comparator(s)
3. Acceptable placebo response, defined as a Drug Liking VAS response between 40 to 60 inclusive
4. Demonstrate responses to active comparator(s) which are consistent with discrimination relative to placebo on all pharmacodynamic measures
5. Able to tolerate the dose of active comparator(s) as judged by the investigator or designee based on available safety data
6. Able to insufflate the entire dose of active comparator(s) (≥95%). Negligible amounts of study drug remaining in containers are acceptable.
7. Demonstrate general behavior suggestive that the subject could successfully complete the study

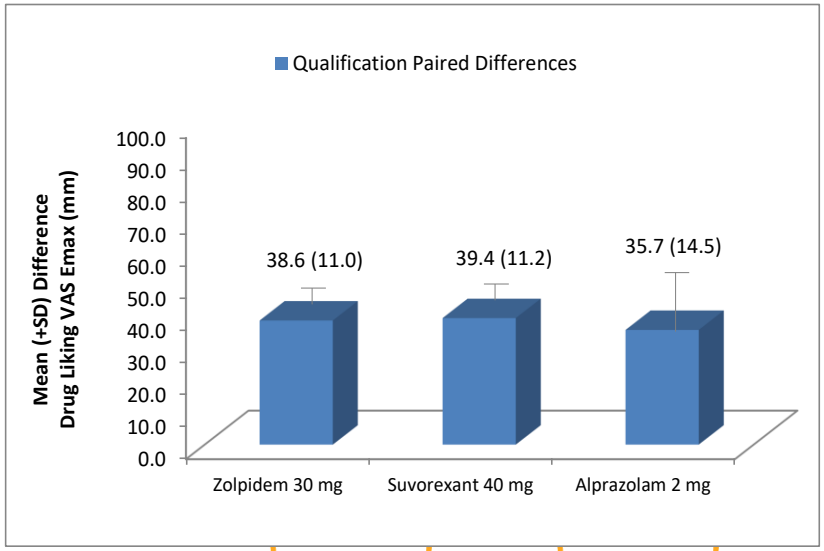
Qualification vs. Treatment Phases: Drug Liking VAS E_{max}

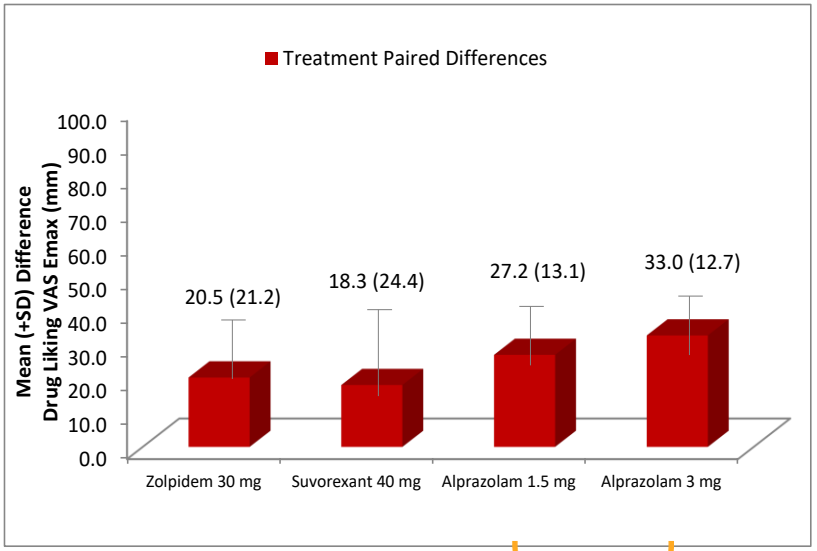
Figure 1. Drug Liking VAS E_{max} for Qualification vs. Treatment Phases Across Various Drug Classes (Positive Controls)



Drug Liking VAS E_{max} - Complex vs Simple Qualification

Figure 2. Paired Differences of Positive Control vs. Placebo for Drug Liking E_{max} for Complex (2 drug) vs Simple (1 drug) Qualification





Complex – 2 Drug Qualification

Simple – 1 Drug Qualification

2 mg alprazolam Qualification

Responder Analysis in the Treatment Phase – Oral HAP Studies

	Paired Differences			
Response Category	**<0**	**0**	**(0,15]**	**>15**
Phentermine 45 mg n=37	**2 (5.4)**	**8 (21.6)**	**6 (16.2)**	21 (56.8)
Phentermine HCl 60 mg n=38	**0**	**8 (21.1)**	**8 (21.1)**	22 (57.9)
Zolpidem 30 mg n=32	**5 (15.6)**	**1 (3.1)**	**6 (18.8)**	20 (62.5)
Suvorexant 40 mg n=32	**5 (15.6)**	**3 (9.4)**	**5 (15.6)**	19 (59.4)
Alprazolam 1.5 mg n=39	**0**	**0**	**5 (12.8)**	34 (87.2)
Alprazolam 3 mg n=39	**0**	**0**	**4 (10.3)**	35 (89.7)
Lorazepam 2 mg n=34	**5 (14.7)**	**1 (2.9)**	**8 (23.5)**	20 (58.8)
Lorazepam 4 mg n=34	**3 (8.8)**	**0**	**7 (20.6)**	24 (70.6)
Marinol 10 mg n=33	**5 (15.2)**	**2 (6.1)**	**5 (15.2)**	21 (63.6)
Marinol 30 mg n=33	**0**	**1 (3.0)**	**4 (12.1)**	28 (84.8)

Inferential Analysis in the Treatment Phase: Oral HAP Study

				$H_o: \mu_C - \mu_P \leq 15$ vs. $H_a: \mu_C - \mu_P > 15$	$H_o: \mu_C - \mu_P \leq 11$ vs. $H_a: \mu_C - \mu_P > 11$
PD Endpoint	Contrast	Mean (SE)	1-sided 95% Confidence Interval	P-value	P-value
Drug Liking VAS E_{max}	Zolpidem 30 mg vs. Placebo	20.2 (3.65)	(14.1 - Infinity)	0.079	0.006
	Suvorexant 40 mg vs. Placebo	18.2 (3.65)	(12.2 - Infinity)	0.190	0.025

Responder Analysis in the Treatment Phase – Intranasal ADF Studies

	Paired Differences			
Response Category	**<0**	**0**	**(0,15]**	**>15**
Stimulant 40 mg n=37	2 (5.4)	2 (5.4)	11 (29.7)	22 (59.5)
Oxycodone HCl 15 mg n=42	0	3 (7.1)	4 (9.5)	35 (83.3)
Oxycodone HCl 30 mg n=37	0	1 (2.7)	0	36 (97.3)
Hydrocodone 50 mg n=21	0	2 (9.5)	1 (4.8)	18 (85.7)

Inferential Analysis in the Treatment Phase: Intranasal ADF Study

			Ho: μC - μP ≤ 15 versus Ha: μC - μP > 15		Ho: μC - μP = 0 versus Ha: μC - μP ≠ 0	
PD Endpoint	Contrast	Median (Q1 – Q3)	1-sided 97.5% Confidence Interval	P-value	2-sided 95% Confidence Interval	P-value
Drug Liking VAS E_{max}	Stimulant 40 mg vs. Placebo	23.0 (11.0 - 37.0)	(13.0 - Infinity)	0.088	(13.0 - 32.0)	<.0001

Other Considerations

- “All the margins should be pre-specified and justified in the protocol”.

  page 29 of 2017 FDA Guidance

  - May not be possible for molecules with novel or partially elucidated mechanisms of action

- “For each hypothesis, the statistical significance of the test should be achieved on all doses”, page 29 of 2017 FDA Guidance

  - HAL studies aim to meaningfully estimate the likelihood that a new drug might be used for recreational rather than medical purposes
  - Reaching statistical significance for all the planned analyses is not a pre-requisite for determining potential for abuse

Summary Points

- Data presented is limited to 10 studies.
- Larger margins may lead to failed studies.
- Drugs with known abuse potential may appear as false negatives, that is, positive control does not separate from placebo.
- Increasing the complexity of Qualification Phase does not reduce variability in Treatment Phase.
- Analyses for all PD measures should be performed to fully understand the profile of a yet to be tested drug for which no exhaustive set of apriori hypotheses can be reasonably developed.
- Eliminating or reducing margins is required to conduct adequate evaluations.