CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

207975Orig1s000

STATISTICAL REVIEW(S)

Statistical Team Leader Memorandum

Submission: NDA 207975/000
Product: Vantrela (hydrocodone bitartrate) extended-release tablets
Sponsor: Teva Pharmaceuticals, Inc.
Indication: Management of pain severe enough to require daily, around-the-clock, long-term opioid treatment and for which alternative treatment options are inadequate
Medical Division: Division of Anesthesia, Analgesia, and Addiction Products (DAAAP)

Reference: Statistical Review and Evaluation dated September 11, 2015

The purpose of this memorandum is to clarify the conclusions in the primary reviewer's evaluations of this original NDA submission, and to provide more details on the interpretation of the sensitivity analyses.

The primary review focused on the positive phase 3 trial 3103 and concluded "the amount of missing data in study 3103 coupled with the marginal effect in those with week 12 data does not lead to robust evidence in favor of Vantrela providing greater relief of low back pain than placebo". This conclusion, albeit an accurate description of the results, may be ambiguous.

The primary efficacy endpoint of trial 3103 was change from baseline to week 12 in the weekly average of worst pain intensity (WPI). The primary analysis was ANCOVA model with baseline WPI, randomized treatment, opioid status, and center as covariates. The intent-to-treat analysis population, defined as all randomized patients, was used for the primary efficacy analysis.

The applicant performed multiple imputation on the week 12 missing data for the primary analysis. The imputation model included randomized treatment, opioid status, baseline and post-baseline WPI values while subjects in the active-drug treatment group who discontinued study drug because of an adverse event, were treated as if they were in the placebo group and their missing data were imputed based on the observed placebo subjects' data. The pre-specified primary analysis showed statistical significance on the primary endpoint (p-value = 0.0012). One of the flaws of this imputation method is that the observed data in placebo subjects are not representative of all placebo data. Also as discussed in the primary review, this missing data imputation may not be an ideal analysis to describe a *de facto* estimand that are usually aimed for a clinical trial. However, the results of such an analysis are of potential to support treatment efficacy as this imputation method addresses to the usual concern about assigning favorable values to discontinuation due to intolerance of the treatment. The primary review included a sensitivity analysis where the retrieved data on the subjects discontinued treatment were incorporated in the primary analysis and the results were still statistically significant (p-value = 0.0068). The primary review engaged an extensive tipping point analysis to further investigate the impact of missing data on the primary analysis results. Note that the confidence interval bounds were used instead of p-values for statistical significance determination in Table 11 of the primary review. This tipping point analysis provides a comprehensive understanding of the magnitude and degree of the missing data. The results of this analysis should be taken into consideration in the treatment efficacy determination. However, the approval/non-approval

decision of the investigative treatment should be based on the totality of data, and not solely on the results of the tipping point analysis.

After a thorough review, we requested an additional sensitivity analysis with all subjects in the active-drug treatment group who discontinued study drug treated as if they were in the placebo group and their missing data were imputed based on the observed placebo subjects' data regardless of the discontinuation reasons. This sensitivity analysis cannot resolve the interpretation flaws of the primary analysis. However, it renders a different perspective of the data and the results provide supportive evidence of the treatment efficacy. The results of this sensitivity analysis reported by the applicant along with the results from the primary review are presented in Table 1 below.

Table 1. Analysis of change in WPI from Baseline to Week 12

	Vantrela ER	**Placebo**
Applicant's primary analysis		
N*	152	133
Adj. mean change from baseline	0.1	0.7
HER – Placebo	-0.6	
(95% CI)	(-1.00, -0.25)	
p-value	0.0012	
Sensitivity analysis 2 (preferred FDA analysis)		
N*	161	145
Adj. mean change from baseline	0.1	0.7
HER – Placebo	-0.5	
(95% CI)	(-0.90, -0.14)	
p-value	0.0068	
Sensitivity analysis 3 (requested)		
N	191	179
Adj. mean change from baseline	0.1	0.7
HER – Placebo	-0.6	
(95% CI)	(-0.97, -0.24)	
p-value	0.001	

** Number of subjects with week 12 data included in the analysis; Analysis based on the 1000 imputed datasets*

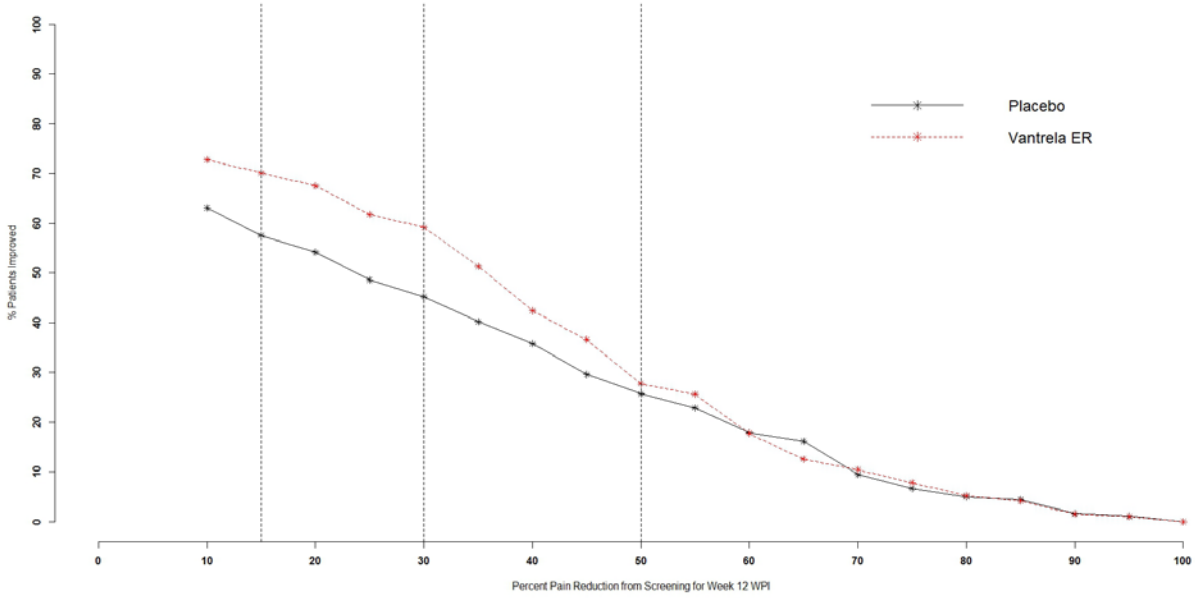
The treatment efficacy can be further supported by the ancillary responder analysis results on the primary endpoint. All patients missing Week 12 values are treated as non-responders in this analysis. The results of this analysis are reproduced by this reviewer and displayed in Table 2 and Figure 1 below. A consistent and nominally significant effect over a relevant range can be

observed, although this analysis is not formally controlled for multiplicity and should only be viewed descriptively.

Table 2. Improvement by 15%, 30% and 50% in WPI from Screening to Week 12

Threshold	Vantrela ER (n, %)	Placebo (n, %)
N	191	179
15%	134 (70%)	103 (58%)
30%	113 (59%)	81 (45%)
50%	53 (28%)	46 (26%)

Figure 1. Percentage Improvement in WPI from Screening to Week 12



In conclusion, the results of the primary analysis along with the sensitivity and ancillary analyses have provided sufficient evidence on the efficacy of Vantrela in moderate to severe chronic low back pain management, as measured by the change from baseline in the weekly average WPI at week 12.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

FREDA COONER
09/28/2015

THOMAS J PERMUTT
09/28/2015
I concur.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA #:	207975
Drug Name:	Vantrela (hydrocodone bitartrate) extended-release tablets
Indication:	Management of pain severe enough to require daily, around-the-clock, long-term opioid treatment and for which alternative treatment options are inadequate
Applicants:	Teva Pharmaceuticals, Inc.
Date(s):	Stamp date: 12/23/2014 Review due date: 9/18/2015
Review Priority:	Standard
Biometrics Division:	II
Statistical Reviewer:	Bradley W. McEvoy, DrPH
Concurring Reviewers:	Freda Cooner, PhD
Medical Division:	Anesthesia, Analgesia, and Addiction Products
Clinical Team:	Robert A. Levin, MD, Medical Reviewer John Feeney, MD, Clinical Team Lead
Project Manager:	Kimberly Compton

Keywords: NDA review, missing data

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

Teva Pharmaceuticals, Inc. submitted a new drug application for Vantrela (hydrocodone bitartrate) extended-release, abuse deterrent, oral tablet in 15, 30, 45, 60 and 90 mg strengths. The application is supported by one positive pivotal phase 3 trial (protocol 3103) and one failed pivotal trial (protocol 3079). The primary emphasis of this review is on the positive trial, which had a multicenter, randomized, double-blind, placebo-controlled, 12 week randomized withdrawal design with an open-label titration period that lasted up to 6 weeks to assess the efficacy and safety of Vantrela in patients with moderate to severe chronic low back pain who require continuous opioid treatment for an extended period of time.

Based on the applicant's prespecified analysis, placebo had a statistically significantly greater increase than Vantrela in mean worst pain intensity (WPI). The estimated WPI change was 0.6 units smaller for Vantrela than placebo with 95% CI (0.25, 1.00). The primary analysis was found to have notable interpretation issues since it attempts to describe a hypothetical rather than real-world drug effect. Specifically, the primary analysis attempts to describe the effect assuming all subjects could adhere to study treatment for 12 weeks, contrary to the fact that some subjects could not adhere. In addition to the inference not being consistent with the design of trial 3103, there is also an issue that analysis does not meet statutory requirement of being able to establish "the drug will have the effect it purports or is represented to have."

In total, 625 were enrolled in the trial, with 371 subjects entering the double-blind study period, with 180 randomized to placebo and 191 to Vantrela. Two hundred ninety seven subjects (297, 80%) completed the 12-week double-blind period, with a slightly greater number completing in the Vantrela group than in the placebo group (82% vs. 78%). The percentage of treatment completers was 77% for Vantrela, which was slightly greater than 72% for placebo.

The number of subjects with week 12 WPI data was greater in the Vantrela group than placebo (84% vs. 81%). Week 12 data measured after subjects discontinued study treatment were excluded from the applicant's primary analysis. Therefore, the number of subjects with week 12 data that contributed to the analysis was less than the number of subjects with week 12 data. In total, 80% in the Vantrela group and 74% in placebo had week 12 data and did not discontinue study drug.

In a supportive analysis that included data subjects discontinued study treatment, Vantrela provided a greater relief of low back pain in placebo at week 12. However, a systematic evaluation into the potential impact of missing data on the difference in week 12 WPI change regardless of treatment adherence revealed the overall results could be susceptible to violations in assumptions about the missing data. There were scenarios investigated that could possibly describe the experience of those with missing data (e.g., equal WPI change in the treatment groups among those with missing data) that caused the results to no longer be statistically significant (See Section 3.2.1.4.1.)

A total of 26 randomized subjects were potentially un-blinded during the double-blind treatment period of study 3103 based on the titration scheme in the optional long-term safety roll-over

study, study 3104. To investigate the impact of potential un-blinding the applicant repeated the primary analysis excluding the 26 subjects whose treatment may have been unblinded. Results from this analysis were consistent with the primary analysis, suggesting that the conclusion was not impacted by the potential un-blinding.

Study 3079 failed to demonstrate a statistically significant improvement on the primary study endpoint, average pain intensity. Based on the planned analysis for the secondary endpoint WPI, the difference between Vantrela and placebo was statistically significant at the nominal 5% significance level. This finding does not however provide an independent substantiation of experimental results from study 3103 because the result on WPI from 3079 was hypothesis generating and susceptible to various biases, including an inflation of type I error and random high bias.

In conclusion, the amount of missing data in study 3103 coupled with the marginal effect in those with week 12 data does not lead to robust evidence in favor of Vantrela providing greater relief of low back pain than placebo, as measured by the change from baseline in the weekly average WPI at week 12.

2 INTRODUCTION

2.1 Overview

On 12/23/2014, Teva Pharmaceuticals, Inc. submitted a 505(b)(2) application for Vantrela (hydrocodone bitartrate) extended-release, abuse deterrent, oral tablet in 15, 30, 45, 60 and 90 mg strengths. The application is supported two pivotal phase 3 trials – one failed (protocol 3079) and one positive (protocol 3103). Primary emphasis in this review is on the positive trial entitled, "A 12-week, randomized, double-blind, placebo-controlled, randomized-withdrawal study to evaluate the efficacy and safety of hydrocodone bitartrate extended-release tablets (CEP-33237) at 30 to 90 mg every 12 hours for relief of moderate or severe pain in patients with chronic low back pain who require opioid treatment for an extended period of time."

2.1.1 Class and Indication

Vantrela (hydrocodone bitartrate) extended release is an abuse deterrent semi-synthetic opioid. In November 2014, FDA approved the first extended release hydrocodone bitartrate with abuse deterrent properties (Hysingla, NDA 206627). The proposed indication for Vantrela is for management of pain severe enough to require daily, around-the-clock, long-term opioid treatment and for which alternative treatment options are inadequate.

2.1.2 History of Drug Development

The applicant's hydrocodone bitartrate extended release (HER) was developed under IND 105587. The application was submitted under the 505(b)(2) pathway and relies on FDA's previous findings of safety and efficacy for the immediate-release hydrocodone component of Vicoprofen (NDA 20716). On 7/21/2015 the applicant submitted a request that the application be converted to a 505(b)(1) application. The applicant obtained the right of reference to the Vicoprofen NDA from AbbVie, Inc. In an email issued on 2/9/2015, FDA told the applicant that they would not impose additional requirements for a 505(b)(1) application beyond what would have been required if they submitted a 505(b)(2) application. This advice, however, contrasts to the advice given at the Type C Guidance meeting held on 1/15/2015.

At the Type B meeting held on 7/14/2010, FDA conveyed to the applicant that for a 505(b)(2) submission a single adequate and well-designed clinical trial would be sufficient to support the indication. The applicant designed protocol 3079 as the pivotal study for the application, which failed to meet its primary objective. The applicant designed protocol 3103 as a second pivotal study, which differed in several important study design features from protocol 3079 (See next section for details). At the Type A meeting on 9/6/2012 FDA stated protocol 3103 could support the indication, although data from study 3079 would be reviewed as part of the application. At the 1/15/2015 Type C meeting, FDA stated that, if the applicant wanted to pursue the 505(b)(1) pathway, it would be a review issue of whether study 3079 could fulfill the requirement for a second study to support a finding for efficacy.

2.1.3 Specific Studies Reviewed

Design features of the two trials reviewed in this review are summarized in Table 1. The trials differed significantly, including the study population, the primary efficacy endpoint, and doses investigated. The emphasis of this review is on trial 3103 as trial 3079 did not meet its primary objective.

Table 1. Summary of study designs.

Trial	3079	3103
Study periods	-Screening period (7-14 days) -Open-label titration period (up to 6 weeks) -Double-blind treatment period (12 weeks)	-Screening period (7-14 days) -Open-label titration period (up to 6 weeks) -Double-blind treatment period (12 weeks)
Design*	R, DB, PC, PG, RW	R, DB, PC, PG, RW
Population	Osteoarthritis or low back pain	Moderate to severe chronic low back pain
Primary endpoint	Change from baseline to week 12 in the weekly API	Change from baseline to week 12 in the weekly WPI
Doses investigated	15, 30, 45, 60, 90 mg	30, 45, 60, 90 mg
Sample size (N randomized)	HER: 146 Placebo: 148	HER: 191 Placebo: 180

* Double-blind treatment period; R – randomized; DB – double-blind; PC – placebo-controlled; PG – parallel-group; RW – randomized withdraw; API –average pain intensity; WPI –worst pain intensity; HER –hydrocodone bitartrate extended release

2.2 Data Sources

The data and final study report were submitted electronically as an eCTD submission. The submission is archived at the following link: \\CDSESUB1\evsprod\NDA207975\207975.enx.

The following documents were used to support this review:

Clinical study report (trials 3079, 3103)
Study protocol (trials 3079, 3103)
Statistical analysis plan (trial 3103)
Response to 8/24/2015, 2015 FDA information request
Response to 6/15/2015, 2015 FDA information request
Statistical review of the analysis plan for trial 3103 (DARRTS: 10/31/2013)

All results presented in this review were derived from the submitted datasets by this reviewer. All tables and figures in this review were created by this reviewer unless noted otherwise.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The submission quality was found to be reasonably high.

3.2 Evaluation of Efficacy

This Section is organized as follows. Section 3.2.1 provides a thorough evaluation of efficacy data from Trial 3103. Section 3.2.2 provides an overview of Trial 3079, including differences with Trial 3103 and results overall by opioid status for worst pain intensity (WPI) (secondary endpoint; primary endpoint in Trial 3103).

3.2.1 Trial 3103

3.2.1.1 Study Design and Endpoints

Study design: Study 3103 was a phase 3, multicenter, randomized, double-blind, placebo-controlled, randomized withdrawal trial to assess the efficacy and safety of HER tablets in patients with moderate to severe chronic low back pain who require continuous opioid treatment for an extended period of time. Males and females between the ages 18 and 80 with moderate to severe chronic low back pain for at least 3 months that were opioid naïve or experienced were eligible for the study if they were not taking at least 135 mg/day of oxycodone during the 14 days prior to screening and did not have radicular (nerve compression) pain or another type of neuropathic pain. Opioid naïve was defined as patients who were taking tramadol or less than 10 mg/day of oxycodone for 14 days before screening. Opioid experienced was defined as patients who were taking less than 135 mg/day and at least 10 mg/day of oxycodone for the 14 days before screening.

The trial consisted of a 7 to 14 day screening period, an open-label titration period that lasted up to 6 weeks, and a 12 week double-blind treatment period. Informed consent and eligibility for the open-label period were determined during the screening period.

During the open-label titration period, the dose of HER that produces stable pain relief without unacceptable adverse events (AEs) and without the patient exceeding the allowable dose of rescue medication was to be identified. Stable pain relief was defined as an average pain intensity (API) score over the past 24 hours of 4 or less and a WPI score of 6 or less on the NRS-11 for either 4 consecutive days or 4 out of 7 consecutive days without unacceptable AEs, while the patient is maintained on the same dose of study drug for up to 7 days. WPI and API were recorded daily in an electronic diary. Opioid-naïve patients were initiated at the 15 mg dose. Opioid experienced patients were initiated at the dose that is approximately half the screening dose. Dose titration occurred every 3 to 7 days until the criterion for a successful dose was achieved. Patients were eligible for the double-blind treatment period once they reached the

successful dose of study drug and that dose was at least 30 mg every 12 hours and no more than 90 mg every 12 hours.

At visit 7 (baseline), eligible subjects were randomized by site and opioid status in a 1:1 ratio to HER or matching placebo every 12 hours. Patients assigned to placebo had a two-week HER tapering period to minimize the withdrawal effects. The doses given every 12 hours during this period for placebo is shown in Table 2. The study blind was maintained during the tapering period. Starting at the third week (visit 9) and for the remainder of the double-blind period subjects received either HER or matching placebo.

Table 2. HER tapering schedule during the double-blind treatment period for the placebo group based on successful HER dose identified from the open-label titration period

	Successful HER dose identified from the open-label titration period			
	30 mg	45 mg	60 mg	90 mg
1st week	HER: 15 mg (×1) PLA: 30 mg (×1)	HER: 30 mg (×1) PLA: 45 mg (×1)	HER: 30 mg (×1) PLA: 30 mg (×1)	HER: 45 mg (×1) PLA: 45 mg (×1)
2nd week	HER: 15 mg (×1) PLA: 30 mg (×1)	HER: 15 mg (×1) PLA: 50 mg (×1)	HER: 15 mg (×1) PLA: 30 mg (×2)	HER: 15 mg (×1) PLA: 45 mg (×2)

HER – hydrocodone bitartrate ER; PLA – placebo

Rescue medication was permitted throughout the double-blind treatment period. The amount permitted (irrespective of tapering period) was as follows: hydrocodone (5 mg)/acetaminophen (325 mg) tablets, 1 to 2 tablets every 4 to 6 hours (as needed), not to exceed a total of 12 tablets or a total dosage of 60 mg hydrocodone or 3900 acetaminophen per day. Patients were to record pain intensity before taking rescue medication.

Patients that discontinued study drug and maintained consent were to be followed according to their regular study visits as specified in the protocol and have a final on-treatment visit. Patients that withdrew from the study were to attend an early termination visit.

Primary efficacy endpoint: Change from baseline to week 12 in the weekly average of WPI. Weekly average of WPI is calculated using WPI scores from the previous 7 days for each study visit.

Sample size: The applicant calculated that a sample size of 170 patients per treatment arm would provide 90% power at a two-sided 5% alpha level to detect a 0.7 difference between HER and placebo on WPI change from baseline to week 12, assuming a standard deviation of at least 2.0.

Secondary efficacy endpoints:

- Change from baseline in weekly average of daily API scores at week 12, based on an NRS-11 form in an electronic diary.
- Time to loss of efficacy, defined as discontinuation of study drug for lack of efficacy, or the start of excessive rescue medication use while on study drug. Excessive rescue medication was defined as 10 or days of rescue medication usage in any 14 consecutive days at a total of 15 mg (hydrocodone-equivalent) or higher each day during the post 2-week tapering period of the double-blind treatment period.

- Percentage of patients with both a 30% or greater increased in API from baseline to week 12 and an API score of 5 or higher at week 12
- Change from baseline to the final on-treatment visit in the Roland Morris Disability Questionnaire (RMDQ). RMDQ is a patient-rated, 24-question evaluation that attempts to assess acute disability with low back pain. Each question is answered with either a YES or NO response, with a YES response assigned a value of 1 and a NO response assigned a value of 0. Scores on RMDQ are based on the sum of the responses to the 24 questions, with higher scores indicating greater disability.

3.2.1.2 Statistical Methodologies

Applicant's primary efficacy analysis: Change in WPI from baseline to week 12 was analyzed using an ANCOVA model with baseline WPI, randomized treatment, opioid status, and center as covariates. WPI values measured after discontinuation of study drug were excluded from the analysis. These data along with missing week 12 WPI data were imputed based on on-treatment data using multiple imputation.

Imputed data were obtained from an imputation model that included assigned treatment, opioid status, baseline and post-baseline WPI values. The applicant's results were based on 5 imputed datasets. To minimize the randomness in the estimates that are associated with having a small number of imputed datasets, analyses presented in this review are based on 1000 simulated datasets. This difference is likely to cause differences between the results in my review and the applicant's study report. Patients assigned to HER and discontinued study drug due to an adverse event were imputed as if they were assigned to placebo; this was achieved by, for the imputation analysis only, recoding their assigned treatment as placebo, not HER.

Primary analysis population: Intention-to-treat analysis set, defined as all randomized patients.

Estimands: It appears the goal of the pre-specified primary analysis was to describe de Jure estimand or the effect on week 12 WPI change if all subjects could adhere to study drug. This estimand represents a hypothetical drug effect as it is not supported by the study design (i.e., subjects were not forced to stay on study drug for the entire study duration). Because of this, from a regulatory perspective, I do not believe the analysis meets the statutory requirement of establishing that "the drug *will* have the effect it purports or is represented to have" (Food, Drug, and Cosmetic Act) because some patients will not be able to adhere to study drug in real-world clinical practice. To these concerns, FDA conveyed in an email sent on 12/14/2012 that the causal estimand that is clinically relevant for the desired indication.

It should be noted that not all of the applicant's sensitivity analyses could evaluate the robustness of findings to deviations in modeling assumptions in the primary analysis. Some analyses are characterized by either the addition of or exclusion of endpoint data relative to that included in the primary analysis. Such analyses are most appropriately viewed as attempting to describe alternative drug-effects or estimands. Only one of these analyses estimates a real-world drug effects, which is of interest.

The applicant's second sensitivity analysis attempts to describe a real-world effect by utilizing measurements after discontinuation of study drug. This analysis attempts to estimate the difference in WPI at week 12 regardless of treatment adherence, which corresponds to the intention-to-treat (ITT) estimand. However, I question whether the applicant's analysis reliably estimates this effect due to concerns with the imputation approach because it does not account for the distribution of an important effect modifier (i.e., treatment) being differentially distributed among those with and without data. In spite of this, it is possible to evaluate whether the conclusion can be adversely impacted by violations in assumptions regarding missing data by means of a tipping-point analysis, which is included in this review.

This review will consider another analysis that describes a real-world drug effect by considering a composite variable defined by treatment adherence up to week 12, and change in week 12 WPI. Details on this analysis are provided below.

The applicant's planned sensitivity analyses for the primary efficacy endpoint: The sensitivity analyses used the same statistical model and imputation strategy for the primary analysis, but differed in what data were used. The implication of using including/excluding data was discussed above. The different types of data used in the four sensitivity analyses are:

- *Sensitivity analysis 1*: Excluded data after the first dose of excessive rescue. According to the applicant, this analysis attempts to assess the impact of excessive rescue.
- *Sensitivity analysis 2*: Included data after discontinuation of study drug.
- *Sensitivity analysis 3*: The calculation of average WPI used, for a given day, the largest WPI values measured during the day. This one is not inclusion/exclusion data.
- *Sensitivity analysis 4*: Same data censoring as the primary analysis, but excludes data from 26 subjects that rolled into the long-term safety study 3104.

FDA analysis of a composite of adherence and WPI: A response on the composite is defined as not having stopped study treatment prior to week 12, and a week 12 change in WPI from screening $\leq$ X% of screening. Note that the WPI at screening was used (instead of WPI at baseline) in the composite definition to be consistent with information presented in the Hysingla labeling. Thresholds of 15%, 30% and 50% will be formally investigated. Note in this approach the 64 subjects with missing week 12 data are classified as non-responders because none of them completed the 12 week treatment period. Other WPI thresholds are explored, with results presented graphically.

FDA evaluation of the potential impact of missing data on WPI change at week 12: A tipping-point analysis was used to explore whether missing data could have adversely impacted findings from the analysis of week 12 WPI change regardless of treatment adherence. The goal of a tipping-point analysis is to explore how the overall results may be impacted by different assumptions about the magnitude of week 12 WPI change from baseline for those with missing week 12 data, which are sensitivity parameters (defined for each treatment group) that are systematically varied. Of interest is to identify the region in the sensitivity parameter space that leads the results to no longer be statistically significant. Concerns with the potential impact of missing data would be either lessened or heightened depending on whether the values of the

sensitivity parameters that tip the results could plausibly describe what may have happened at week 12 in those with missing week 12 data.

The tipping-point analysis was implemented as follows:

1. 100 imputed datasets were created using multiple imputation (MI), where the missing values were imputed using information from the subjects with week 12 WPI values. The imputation was done within treatment group, with baseline WPI as a covariate.
2. For each subject with a missing week 12 value:
    a. The average imputed value was computed
    b. The average imputed value was subtracted from the imputed data, resulting in the imputed data for a subject having mean zero
    c. A constant Δ_{HER} (Δ_P) was added to the imputed value for HER (Placebo)
3. Results from an ANCOVA model fit to the imputed datasets were combined using Rubin's method (Rubin, D.B., *Multiple Imputation for Nonresponse in Surveys*. J. Wiley & Sons, New York, 1987.). The ANCOVA model included baseline WPI, opioid status, center and treatment as covariates. Treatment effect estimates and limits from the 95% confidence interval (CI) were retained.
4. The above steps were repeated, using different values for Δ_{HER} and Δ_P.

Control of type-I error: The study-wise type-I error was controlled at a two-sided 5% level using a hierarchical testing strategy; the order in which the secondary endpoints were tested is listed in Section 3.2.1.1.

Analysis of secondary endpoints:

- *Change from baseline in weekly average of daily API scores at week 12*: Same statistical approach (i.e., type of measured data excluded, imputation and statistical model) as implemented for primary analysis.
- *Time to loss of efficacy*: Cox proportional hazard model, with the censoring flag set to zero if the patient discontinued due to lack of efficacy or used excessive rescue (defined above), and one otherwise.
- *Both a 30% or greater increased in API from baseline to week 12 and an API score of 5 or higher at week 12*: Logistic regression stratified by center with treatment, opioid status, and baseline API as covariates. Week 12 API measured after treatment discontinuation and missing values were imputed using the same approach as the primary analysis.
- *Roland Morris Disability Questionnaire (RMDQ)* The difference in the change from baseline to the last on-treatment RMDQ value is compared using an ANCOVA model with the following variables as covariates: treatment, center, opioid status, and baseline RMDQ.

3.2.1.3 Patient Disposition, Demographic and Baseline Characteristics

Patient disposition: In total, 625 were enrolled in the trial, with two not being treated during the open-label titration period; two subjects (10397003 and 10405014) withdrew consent before taking any study drug. Of the 623 subjects treated during the open-label titration period, 371 subjects entered the double-blind study period, with 180 randomized to placebo and 191 to HER (Table 3). In total, 297 (80%) subjects completed the 12 week double-blind period, with the percent being slightly greater for the HER group than that for the placebo group (82% vs. 78%). Approximately 75% of subjects completed study drug treatment and the double-blind treatment. The percentage of treatment completers was 77% for HER, which was slightly greater than the 72% for placebo.

The most common reason for discontinuing treatment was an adverse event for HER (5%) and lack of efficacy for placebo (8%). Three percent (3%) in the HER group stopped treatment because of lack of efficacy.

Table 3. Subject disposition (Study 3103)

	Hydrocodone ER n (%†)	Placebo n (%†)
Randomized	191	180
Evaluable for efficacy	191 (100%)	179 (99%)
Completed study	156 (82%)	141 (78%)
Completed study, but not study treatment	9 (5%)	11 (6%)
Withdrawn from study	35 (18%)	39 (22%)
Adverse event	9 (5%)	5 (3%)
Lack of efficacy	4 (2%)	9 (5%)
Non-compliance: study drug administration	2 (1%)	2 (1%)
Non-compliance: study procedures	1 (1%)	3 (2%)
Consent withdrawn	9 (5%)	8 (4%)
Protocol violation	7 (4%)	9 (5%)
Lost to follow-up	2 (1%)	2 (1%)
Pregnancy	0	1 (1%)
Other	1 (1%)	0
Completed study treatment	147 (77%)	130 (72%)
Discontinued study treatment	44 (23%)	50 (28%)
Adverse event	15 (8%)	9 (5%)
Lack of efficacy	5 (3%)	17 (9%)
Non-compliance: study drug administration	2 (1%)	2 (1%)
Non-compliance: study procedures	1 (1%)	2 (1%)
Consent withdrawn	9 (5%)	8 (4%)
Protocol violation	7 (4%)	9 (5%)
Lost to follow-up	2 (1%)	1 (1%)
Pregnancy	0	1 (1%)
Other	3 (2%)	1 (1%)

* All randomized patients; † Percent of randomized subjects

Subject disposition was summarized by opioid status to explore whether there were systematic differences between opioid naïve and opioid experienced subjects (Table 4). Differences between the groups were observed. Among the naïve patients, a similar percentage of patients in the two treatment groups (75%) completed 12 weeks of study treatment. For opioid experienced there were notably more subjects in HER group completed the drug the 12 week double-blind than in placebo group (80% vs. 68%). The greater number of treatment discontinuation in placebo for opioid experienced appears to be driven by lack of efficacy. For the opioid naïve group, there were no notable imbalances for the reason for treatment discontinuation between the treatment groups.

Table 4. Subject disposition by opioid status (Study 3103)

	Experienced		Naïve	
	Hydrocodone n (%†)	**Placebo n (%†)**	**Hydrocodone n (%†)**	**Placebo n (%†)**
Randomized	81	75	110	105
Evaluable for efficacy	81 (100%)	75 (100%)	110 (100%)	104 (99%)
Completed study	68 (84%)	58 (77%)	88 (80%)	83 (79%)
Completed study, but not study treatment	3 (4%)	7 (9%)	6 (5%)	4 (4%)
Withdrawn from study	13 (16%)	17 (23%)	22 (20%)	22 (21%)
Adverse event	2 (2%)	0	7 (6%)	5 (5%)
Lack of efficacy	2 (2%)	6 (8%)	2 (2%)	3 (3%)
Non-compliance: study drug administration	1 (1%)	1 (1%)	1 (1%)	1 (1%)
Non-compliance: study procedures	1 (1%)	3 (4%)	0	0
Consent withdrawn	3 (4%)	1 (1%)	6 (5%)	7 (7%)
Protocol violation	4 (5%)	5 (7%)	3 (3%)	4 (4%)
Lost to follow-up	0	1 (1%)	0	0
Pregnancy	0	0	0	1 (1%)
Other	0	0	1 (1%)	0
Completed study treatment	65 (80%)	51 (68%)	82 (75%)	79 (75%)
Discontinued study treatment	16 (20%)	24 (32%)	28 (25%)	26 (25%)
Adverse event	3 (4%)	2 (3%)	12	7 (7%)
Lack of efficacy	3 (4%)	13 (17%)	2 (2%)	4 (4%)
Non-compliance: study drug administration	1 (1%)	1 (1%)	1 (1%)	1 (1%)
Non-compliance: study procedures	1 (1%)	2 (3%)	0	0
Consent withdrawn	3 (4%)	1 (1%)	6 (5%)	7 (7%)
Protocol violation	4 (5%)	5 (7%)	3 (3%)	4 (4%)
Lost to follow-up	0	0	2 (2%)	1 (1%)
Pregnancy	0	0	0	1 (1%)
Other	1 (1%)	0	2 (2%)	1 (1%)

* All randomized patients; † Percent of randomized subjects

Missing data and treatment adherence: Eighteen percent (18%) of subjects did not have a week 12 WPI value available (Table 5). The amount of missing data was slightly greater in placebo than in HER group (19% vs. 16%), with imbalance being associated with opioid status (Table 6). For opioid experienced, placebo had more missing data (20% vs. 14%); a similar percentage of missing data between the treatment groups was observed in the opioid naïve group.

A total of 21 subjects had a week 12 measurement and were no longer receiving study drug at the time. Data from these subjects was not included in the primary analysis since the primary analysis only used data while on-treatment. Therefore, the number of subjects that did not contribute week 12 data was greater than the amount of missing data described above. A total of 20% of subjects in the HER group and 26% in the placebo group did not contribute data in the primary analysis.

Table 5. WPI ascertainment (Study 3103)

	Hydrocodone ER n (%)	Placebo n (%)
Missing week 12 value	30 (16%)	34 (19%)
Week 12 value available	161 (84%)	145 (81%)
Did not discontinue study drug	152 (80%)	133 (74%)
Discontinued study drug	9	12

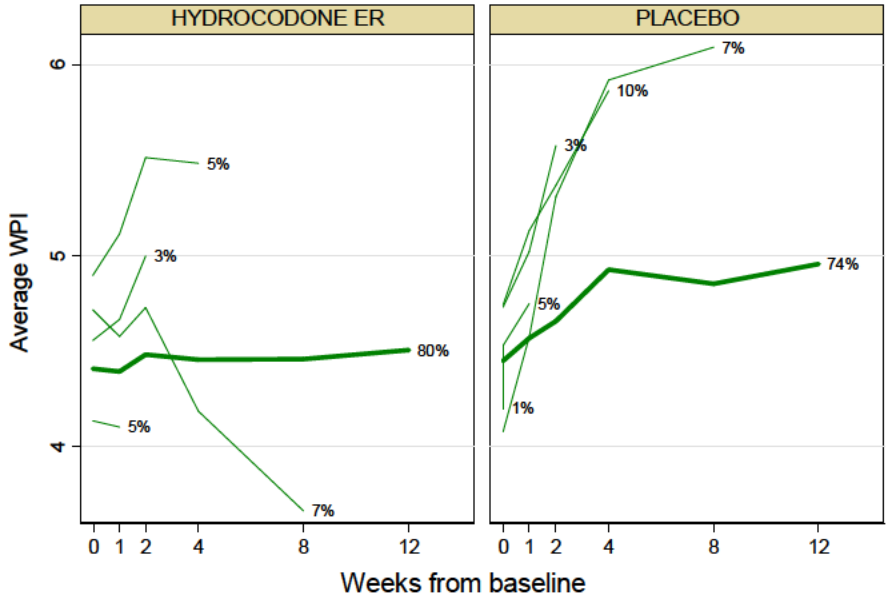
Table 6. WPI ascertainment by opioid status (Study 3103)

	Experienced		Naïve	
	HER n (%)	Placebo n (%)	HER n (%)	Placebo n (%)
Missing week 12 value	11 (14%)	15 (20%)	19 (17%)	19 (18%)
Week 12 value available	70 (86%)	60 (80%)	91 (83%)	85 (82%)
Did not discontinue study drug	67 (83%)	52 (69%)	85 (77%)	81 (78%)
Discontinued study drug	3	8	6	4

On-treatment WPI mean profiles were explored to assess whether differences in treatment discontinuation was possibly associated with the study endpoint (Figure 1). For the placebo group, those that remained on-treatment (thick link) seemed to have more favorable pain scores than those that stopped treatment early. This trend was not observed in the HER group, suggesting there are differences in the experiences for the two treatment groups.

For those that stopped treatment early, the on-treatment WPI (using data from the last visit on treatment) was more favorable for the HER group than placebo. The estimated average WPI was 4.4 for the 39 subjects in HER group and 5.6 for the 46 subjects in placebo group. These numbers must be interpreted cautiously as those in the HER group and placebo group with missing data are not comparable since they are defined by post-baseline events.

Figure 1. WPI mean profile by treatment arm and timing of last on-treatment assessment (off treatment data excluded; Study 3103)



Note: percentages represent number of subjects contributing to the individual profile

Rescue use: During the double-blind treatment period, 136 (71%) of subjects in the HER group and 145 (81%) in the placebo group took rescue medication. It is of concern that these numbers are underestimates of the amount of rescue use during the 12 week follow-up period as subjects that were no longer being followed could not have their rescue use recorded. Because subjects that discontinued study treatment would presumably still need to manage their low back pain, it would be reasonable to investigate the overall amount of non-study medication used during the double-blind period by considering a composite variable defined by rescue use and treatment discontinuation. In total, 151 (79%) subjects in the HER group and 152 (84%) in the placebo group either discontinued study treatment or used rescue medication.

Time-to-excessive rescue use or treatment discontinuation is summarized by treatment groups overall (Figure 2) and by opioid status (Figure 3). There is a dramatic drop at week 2 that coincides with the end of the tapering period. After week 2 the lines for the two groups separate, with a greater number of subjects in placebo using excessive rescue or discontinuing treatment. This difference is associated with opioid status, with the differences being driven by the opioid experienced. The greater separation between curves for opioid experienced is not surprising given there were more subjects that discontinued treatment in placebo than HER (See Table 4).

Figure 2. Time-to-excessive rescue or treatment discontinuation (Study 3103)

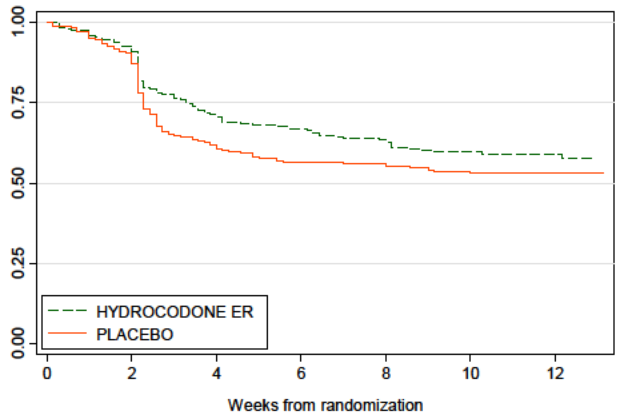
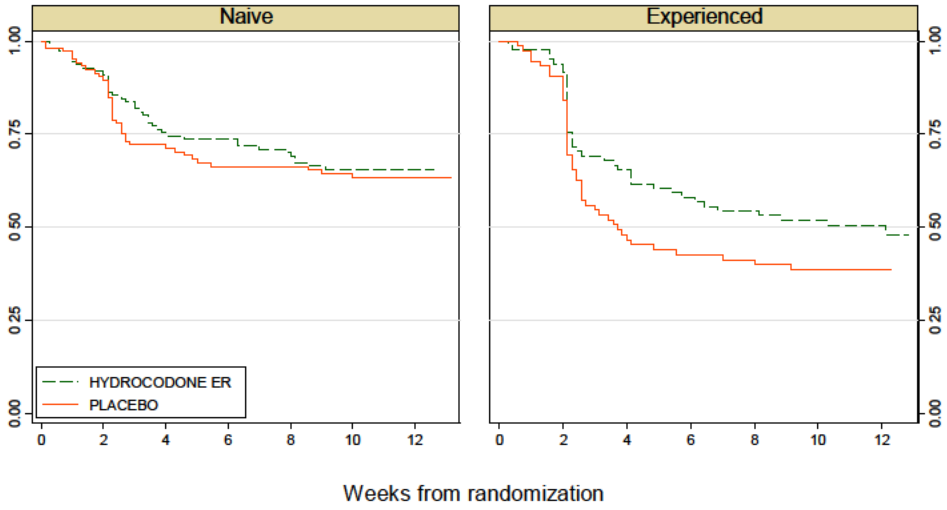


Figure 3. Time-to-excessive rescue or treatment discontinuation by opioid status (Study 3103)



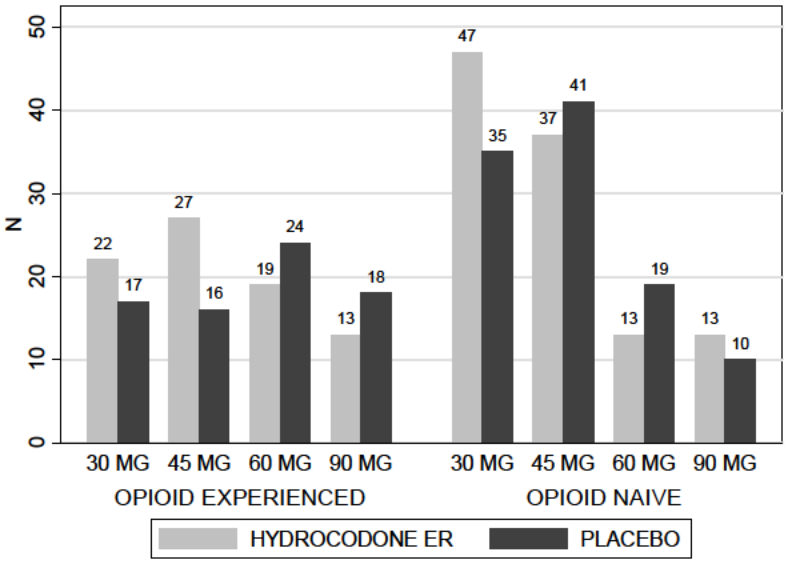
Demographic and baseline characteristics: Demographic and baseline characteristics appeared reasonably similar across the treatment groups (Table 7). The randomized population was 52 years of age on average and there were similar percentages of males and females. The average duration since the diagnosis was approximately 11 years, and subjects have been receiving opioid therapy for approximately 3 years. More patients in the study were opioid naïve, with the opioid naïve patients tending to achieve stable pain relief from the open-label titration period with HER at lower doses than the opioid experienced subjects (Figure 4).

Table 7. Demographic and baseline characteristics by treatment arm (Study 3103)

	HER N=191	Placebo N=180
Age (yrs)		
Mean (SD)	52 (13)	52 (13)
Q2 (Q1, Q3)	52 (42, 62)	53 (43, 61)
≥ 65	37 (19%)	32 (18%)
Gender: Males	94 (49%)	88 (49%)
Race:		
White	133 (70%)	129 (72%)
Black	39 (20%)	41 (23%)
Country: U.S.	191 (100%)	180 (100%)
Opioid status:		
Experienced	81 (42%)	75 (42%)
Naïve	110 (58%)	105 (58%)
Stable dose		
30 mg	69 (36%)	52 (29%)
45 mg	64 (34%)	57 (32%)
60 mg	32 (17%)	43 (24%)
90 mg	26 (14%)	28 (16%)
Duration since CLBP diagnosis† (yrs)		
Mean (SD)	11 (10)	11 (10)
Duration on opioid therapy† (yrs) Mean (SD)	3.5 (4.8)	2.9 (3.8)

† Source: CSR, Table 13; CLBP – chronic low back

Figure 4. Successful HER dose (Ns) from the open-label titration period by opioid status (Study 3103)



Protocol Deviations: A total of 26 randomized subjects were potentially un-blinded to treatment assignment. The issue was that subjects enrolled in 3103 had the option to enroll into study 3104, a long-term safety study, and the titration scheme in 3104 was such that it had the potential to reveal blinded treatment assignment in study 3103. The trial was suspended on 9/17/2013, with the 26 subjects that enrolled into the trial continue with the study without interruption. Enrollment into study 3103 was resumed on 12/2/2013 with a revised titration scheme.

To investigate the impact of potential un-blinding the applicant repeated the primary analysis excluding the 26 subjects whose double-blind treatment may have been unblinded. In this subset, the estimated WPI change was 0.6 units smaller for HER compared to placebo with 95% CI (0.25, 1.03). I was able to replicate the applicant's results.

3.2.1.4 Results and Conclusions

3.2.1.4.1 Primary efficacy endpoint

Primary analysis and preferred FDA analysis: Based on the applicant's prespecified analysis, placebo had a statistically significantly greater mean WPI increase. Compared to placebo, the estimated WPI change was 0.6 units smaller for HER with 95% CI (0.25, 1.00). The study met its primary objective based on the prespecified analysis and testing algorithm.

Results from the preferred FDA analysis (i.e., the applicant's second sensitivity analysis), which included week 12 data from an additional 21 subjects, were in-line with the applicant's primary analysis. The numerical similarity of the results should not be interpreted as providing evidence on the "robustness" of the applicant's primary analysis. The two models quantify different summaries of the intervention effect, with the applicant's analysis having notable limitations (See section 3.2.1.2).

Table 8. Analysis of change in WPI from baseline to week 12

	HER	Placebo
Applicant's primary analysis		
N*	152	133
Adj. mean change from baseline	0.1	0.7
HER – Placebo	-0.6	
(95% CI)	(-1.00, -0.25)	
p-value	0.0012	
Sensitivity analysis 2 (preferred FDA analysis)		
N*	161	145
Adj. mean change from baseline	0.1	0.7
HER – Placebo	-0.5	
(95% CI)	(-0.90, -0.14)	
p-value	0.0068	

** Number of subjects with week 12 data included in the analysis; Analysis based on the 1000 imputed datasets*

Assessment of the missing data imputation from the primary analysis: Imputed values in the primary analysis were investigated to get an understanding of the implicit assumption made in the statistical analysis (Table 9). In total, 39 subjects in HER and 46 in placebo had their week 12 WPI value imputed. For both treatment groups, the week 12 imputed WPI was slightly larger

than that what was observed at the last visit on study drug (HER: 4.6 vs. 4.4; Placebo: 5.8 vs. 5.6), suggesting there is little mediation after stopping treatment.

From Table 9, we can see that the difference in average imputed WPI change between the treatment groups is the same as the difference using the data from the last visit while on study drug (-1.3). This suggests the applicant's primary analysis preserves the treatment difference while on study drug, effectively performing in similar manner as last observation carried forward. The reason for this is likely due to the inclusion of treatment in the imputation model because the difference between groups for those with data is assumed to be the same as those without data. In a response to a statistical information request, the applicant considered this lack of attenuation appropriate because they went back to pain levels before entering the study. I do not agree with the applicant's perspective since you would likely expect greater changes for the HER group, as what was observed in the 9 HER subjects that had stopped treatment and had week 12 measurements. For the 21 subjects that discontinued study drug and had a week 12 WPI measurement, there was no difference in between treatment groups in their change in WPI values at week 12; this suggest there is an attenuation of the treatment effect after subjects discontinue treatment (i.e., patient on placebo get better, where patient on HER tend to get slightly worse).

Table 9. Summary of imputed data from the applicant's primary analysis

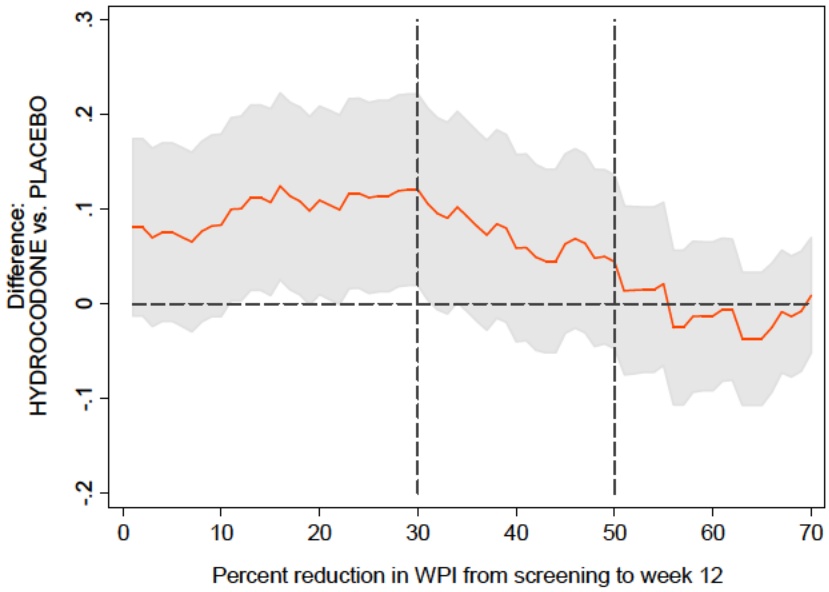
	HER Mean WPI	Placebo Mean WPI	Difference: Δ HER - Δ Placebo
Week 12 WPI imputed for the primary analysis			
N	39	46	
Baseline	4.6	4.5	
Last visit on study drug	4.4	5.6	-1.3
Week 12: Imputed	4.6	5.8	-1.3
Week 12 WPI available but value imputed for the primary analysis			
N	9	12	
Baseline	4.55	4.74	
Last visit on study drug	4.5	5.7	-1.0
Week 12: Imputed	4.6	5.8	-1.1
Week 12: Observed	5.3	5.5	-0.0

FDA composite analysis of treatment adherence: The proportions of subjects that stayed on treatment to week 12 and had a change of at least 15%, 30% and 50% in WPI from screening to week 12 are shown in Table 10 below. For the 15% and 30% threshold the difference between HER and placebo was statistically significant at a nominal 5% significance level. Point-wise difference and confidence intervals for other thresholds are displayed in Figure 5. For almost all thresholds larger than 30% the 95% CI for the difference between groups includes the null value zero.

Table 10. Results from FDA composite of treatment adherence up to week 12 and 15%, 30% and 50% in WPI from screening to week 12 (Study 3103).

Threshold	HER	Placebo	Difference: HER – Placebo (95% CI)
15%	68%	56%	11% (0, 21)
30%	56%	45%	12% (1, 22)
50%	27%	25%	1% (-7, 10)

Figure 5. Point-wise difference (HER – Placebo) and 95% CI for FDA composite of treatment adherence and % change in WPI from screening to week 12 for different thresholds (Study 3103).



Although I consider the above analysis informative, it has a critical limitation. Specifically, the difference in proportions is affected by the effect of the drug on treatment adherence and not solely by effect of the study drug on pain. An important question is whether the subjects that were able to adhere to HER for 12 weeks are better off than they would have been if they were assigned placebo. This effect is very difficult to estimate reliably, but is straightforward (if there are no missing week 12 data in placebo) to obtain a conservative estimate of it. The conservative estimate can be obtained by simply comparing the average response in the HER treatment completers with a subset in placebo with the most favorable WPI change scores.

For this post hoc analysis, missing week 12 WPI in the placebo group was imputed conservatively using a subject's baseline observation. Overall, the 152 (79.6% of randomized subjects) HER treatment completers had an estimated change from baseline to week 12 in WPI of 0.10, corresponding with a slight increase in pain. For the best 142 (79.3% of randomized subjects) subjects in placebo the estimated change in WPI at week 12 using a conservative baseline imputation was -0.21, corresponding to a slight decrease in pain. Overall, the HER group had a greater increase in WPI (difference [HER-placebo] = 0.31; 95% CI: -0.02, 0.65), meaning that the best performing placebo subjects (under a conservative imputation and a conservative comparison method) responded more favorably than the HER treatment completers.

This analysis was repeated in the opioid experienced subjects since the overall results were driven by this subgroup (see Section 4). Overall, the 67 (82.7%) HER treatment completers that are opioid experienced had an estimated increase from baseline to week 12 in WPI of 0.29, which was smaller than 0.39 increase from the best 62 (82.7%) opioid experienced subjects in

the placebo group under conservative baseline imputation. Thus, among opioid experienced subjects, the HER group had a smaller increase WPI than placebo (difference [HER-placebo] = -0.11; 95% CI: -0.64, 0.42), meaning that the best performing placebo subjects (under a conservative imputation and a conservative comparison method) responded slightly less favorable than the HER treatment completers.

FDA tipping point-investigation: The tipping-point investigation suggests that the superiority of HER to placebo with regards to ITT estimand (i.e., the difference in WPI change from baseline to week 12 regardless of treatment adherence) is susceptible to violations in pre-specified assumptions about missing data.

A large number of the scenarios investigated resulted in the difference in average WPI change between treatment groups no longer statistically significant (Table 11). The conclusions were not impacted in scenarios where the HER subjects with missing data had slightly more favorable pain values than placebo subjects with missing data. However, the findings were no longer statistically significant when the WPI change in the subjects with missing data was similar or worse for HER compared to placebo. For example, if the sensitivity parameter for both groups are set to 0.7, coinciding to approximately what was observed for the 21 subjects that stopped treatment and had a week 12 value (Table 9; HER: 5.3 - 4.55 = 0.75; Placebo: 5.5 - 4.74 = 0.76), the estimated treatment effect would no longer be statistically significant. Given the scenarios that caused the finding to no longer be statistically significant could possibly describe the experience of those with missing data, there is concern that the findings relative to the ITT estimand are sensitive to missing data.

Table 11. Missing data sensitivity analysis results: Estimates of the difference in mean WPI change for HER – Placebo (Upper 95% CI) assuming a given WPI change for the group with missing data in the treatment groups

HER sensitivity parameter: Week 12 WPI change from baseline in the group with missing data	Placebo sensitivity parameter: Week 12 WPI change from baseline in the group with missing data															
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5
0	-.3 (.05)	-.32 (.03)	-.34 (.01)	-.36 (-.01)	-.38 (-.03)	-.4 (-.05)	-.42 (-.07)	-.44 (-.09)	-.46 (-.11)	-.48 (-.13)	-.5 (-.15)	-.51 (-.16)	-.53 (-.18)	-.55 (-.2)	-.57 (-.22)	-.59 (-.24)
0.1	-.29 (.06)	-.31 (.04)	-.33 (.02)	-.35 (0)	-.36 (-.02)	-.38 (-.04)	-.4 (-.05)	-.42 (-.07)	-.44 (-.09)	-.46 (-.11)	-.48 (-.13)	-.5 (-.15)	-.52 (-.17)	-.54 (-.19)	-.56 (-.2)	-.58 (-.22)
0.2	-.27 (.08)	-.29 (.06)	-.31 (.04)	-.33 (.02)	-.35 (0)	-.37 (-.02)	-.39 (-.04)	-.41 (-.06)	-.42 (-.08)	-.44 (-.1)	-.46 (-.11)	-.48 (-.13)	-.5 (-.15)	-.52 (-.17)	-.54 (-.19)	-.56 (-.21)
0.3	-.26 (.09)	-.27 (.07)	-.29 (.05)	-.31 (.04)	-.33 (.02)	-.35 (0)	-.37 (-.02)	-.39 (-.04)	-.41 (-.06)	-.43 (-.08)	-.45 (-.1)	-.47 (-.12)	-.49 (-.14)	-.5 (-.15)	-.52 (-.17)	-.54 (-.19)
0.4	-.24 (.11)	-.26 (.09)	-.28 .07)	-.3 (.05)	-.32 (.03)	-.33 (.01)	-.35 (-.01)	-.37 (-.02)	-.39 (-.04)	-.41 (-.06)	-.43 (-.08)	-.45 (-.1)	-.47 (-.12)	-.49 (-.14)	-.51 (-.16)	-.53 (-.17)
0.5	-.22 (.13)	-.24 (.11)	-.26 (.09)	-.28 (.07)	-.3 (.05)	-.32 (.03)	-.34 (.01)	-.36 (-.01)	-.38 (-.03)	-.4 (-.05)	-.41 (-.07)	-.43 (-.08)	-.45 (-.1)	-.47 (-.12)	-.49 (-.14)	-.51 (-.16)
0.6	-.21 (.14)	-.23 (.12)	-.24 (.1)	-.26 (.08)	-.28 (.07)	-.3 (.05)	-.32 (.03)	-.34 (.01)	-.36 (-.01)	-.38 (-.03)	-.4 (-.05)	-.42 (-.07)	-.44 (-.09)	-.46 (-.1)	-.47 (-.12)	-.49 (-.14)
0.7	-.19 (.16)	-.21 (.14)	-.23 (.12)	-.25 (.1)	-.27 (.08)	-.29 (.06)	-.31 (.04)	-.32 (.02)	-.34 (.01)	-.36 (-.01)	-.38 (-.03)	-.4 (-.05)	-.42 (-.07)	-.44 (-.09)	-.46 (-.11)	-.48 (-.12)
0.8	-.17 (.18)	-.19 (.16)	-.21 (.14)	-.23 (.12)	-.25 (.1)	-.27 (.08)	-.29 (.06)	-.31 (.04)	-.33 (.02)	-.35 (0)	-.37 (-.02)	-.38 (-.03)	-.4 (-.05)	-.42 (-.07)	-.44 (-.09)	-.46 (-.11)
0.9	-.16 (.19)	-.18 (.17)	-.2 (.15)	-.22 (.13)	-.23 (.12)	-.25 (.1)	-.27 (.08)	-.29 (.06)	-.31 (.04)	-.33 (.02)	-.35 (0)	-.37 (-.02)	-.39 (-.04)	-.41 (-.05)	-.43 (-.07)	-.45 (-.09)
1.0	-.14 (.21)	-.16 (.19)	-.18 (.17)	-.2 (.15)	-.22 (.13)	-.24 (.11)	-.26 (.09)	-.28 (.07)	-.29 (.06)	-.31 (.04)	-.33 (.02)	-.35 (0)	-.37 (-.02)	-.39 (-.04)	-.41 (-.06)	-.43 (-.07)
1.1	-.13 (.23)	-.14 (.21)	-.16 (.19)	-.18 (.17)	-.2 (.15)	-.22 (.13)	-.24 (.11)	-.26 (.09)	-.28 (.07)	-.3 (.05)	-.32 (.03)	-.34 (.02)	-.36 (0)	-.37 (-.02)	-.39 (-.04)	-.41 (-.06)
1.2	-.11 (.24)	-.13 (.22)	-.15 (.2)	-.17 (.18)	-.19 (.17)	-.2 (.15)	-.22 (.13)	-.24 (.11)	-.26 (.09)	-.28 (.07)	-.3 (.05)	-.32 (.03)	-.34 (.01)	-.36 (0)	-.38 (-.02)	-.4 (-.04)
1.3	-.09 (.26)	-.11 (.24)	-.13 (.22)	-.15 (.2)	-.17 (.18)	-.19 (.16)	-.21 (.14)	-.23 (.13)	-.25 (.11)	-.27 (.09)	-.28 (.07)	-.3 (.05)	-.32 (.03)	-.34 (.01)	-.36 (-.01)	-.38 (-.02)
1.4	-.08 (.28)	-.1 (.26)	-.12 (.24)	-.13 (.22)	-.15 (.2)	-.17 (.18)	-.19 (.16)	-.21 (.14)	-.23 (.12)	-.25 (.1)	-.27 (.09)	-.29 (.07)	-.31 (.05)	-.33 (.03)	-.34 (.01)	-.36 (-.01)
1.5	-.06 (.29)	-.08 (.28)	-.1 (.26)	-.12 (.24)	-.14 (.22)	-.16 (.2)	-.18 (.18)	-.19 (.16)	-.21 (.14)	-.23 (.12)	-.25 (.1)	-.27 (.08)	-.29 (.07)	-.31 (.05)	-.33 (.03)	-.35 (.01)

Shaded boxes correspond to scenarios where the 95% CI for the difference in average WPI change from baseline (HER-Placebo) includes 0.
Source: Results included in applicant's response to the 6/15/2015 FDA information request.

3.2.1.4.2 Secondary endpoints

This section summarizes results from the applicant's analysis of key secondary endpoints. Endpoints are presented in the order they are listed in the hierarchical testing strategy to control the study-wise error at 5%.

API: Based on the applicant's prespecified analysis, the HER group had more favorable change at week 12 compared to placebo that was statistically significant; the estimated difference (HER-placebo) was -0.58 with 95% CI (-0.91, -0.25).

Time to loss of efficacy or start of excessive rescue medication use: A smaller percentage in the HER group had loss of efficacy or used excessive rescue (23% vs. 30%), with the between group difference not being statistically significant based on a time-to-event analysis (hazard ratio = 0.68; 95% CI = 0.45, 1.01; p-value = 0.059). A Kaplan-Meier plot of the time-course was shown previous (Figure 2). The remaining study endpoints cannot be formally tested for statistical significance based on the prespecified testing strategy.

30% increase API from baseline and week 12 API ≥ 5: A greater percentage in the HER group had an increase of at least 30% in API from baseline and API ≥ 5 (19% vs. 13%), with an estimated odds ratio of 0.67 and 95% CI (0.47, 0.96).

RMDQ: The HER group had a smaller reduction from baseline in RMDQ compared to placebo. The estimated difference (HER - placebo) was 0.28 with 95% CI (-0.65, 1.20).

Table 12. Analysis of key secondary efficacy endpoints (Study 3103)

	HER	Placebo
API		
Adj. mean change from baseline	-0.03	0.55
HER – Placebo	-0.58	
(95% CI)	(-0.91, -0.25)	
p-value	< 0.001	
Loss of efficacy or excessive rescue		
n (%)	43 (23%)	54 (30%)
Hazard Ratio (HER/Placebo)	0.68	
(95% CI)	(0.45, 1.01)	
p-value	0.059	
30% increase API from baseline and week 12 API ≥ 5		
%*	13%	19%
Odds ratio (HER/Placebo)	0.67	
(95% CI)	(0.47, 0.96)	
p-value	0.0293	
RMDQ		
Adj. mean change from baseline	-1.29	-1.57
HER – Placebo	0.28	
(95% CI)	(-0.65, 1.20)	
p-value	0.557	

** Based on 152 subjects in the HER group and 133 in the placebo group with week 12 data*

3.2.2 Trial 3079

This section provides an overview of Trial 3079 including results on WPI. Trial 3079 failed to achieve its primary objective.

3.2.2.1 Study Design

Study 3079 was a phase 3, multicenter, randomized, double-blind, placebo-controlled, randomized withdrawal trial to assess the efficacy and safety of HER in patients with moderate to severe pain associated with osteoarthritis (OA) or low back pain who require opioid treatment for an extended period of time. Males and females between the ages 18 and 80 with moderate to severe pain associated with OA or low back pain that were opioid naïve or experienced were eligible for the study.

The trial consisted of a 7 to 14 day screening period, an open-label titration period that lasted up to 6 weeks, and a 12 week double-blind treatment period. During the open-label titration period, the dose of HER that produces stable pain relief without unacceptable AEs was to be identified. Stable pain relief was defined based on API (whereas Trial 3103 defined stable pain relief based on both API and WPI).

At visit 7 (baseline), eligible subjects were randomized by site and opioid treatment status in a 1:1 ratio to HER (15, 30, 45, or 90 mg) or matching placebo every 12 hours. In Trial 3103 there was no 15 mg HER dose. Subjects assigned to placebo entered at two week tapering period to minimize the withdrawal effects. Starting at the third week (visit 9) and for the remainder of the double-blind period subjects received either HER or matching placebo.

The primary efficacy endpoint was change from baseline to week 12 in the API. Change in WPI from baseline to week 12 was a secondary endpoint.

The primary analysis used a similar statistical model including the imputation approach as used in Trial 3103. The main difference was the model did not include study site as a covariate. API data collected after excessive rescue use was excluded from the primary analysis. These data along with missing week 12 data were imputed.

3.2.2.2 Patient disposition, baseline and demographics characteristics, and missing data

A total of 391 subjects were enrolled in the trial, with 294 subjects completing the open-label titration period and were randomized to HER (n = 146) or placebo (n = 148) for the 12-week double-blind treatment period. One subject randomized to placebo was withdrawn from the study by the site investigator prior to receiving double-blind treatment and was not included in the efficacy evaluations. Ninety four (94; 64%) subjects in the HER group completed 12 weeks of

treatment, which was less than the 102 (69%) in the placebo group. Most subjects (27%) achieved stable pain relief with the 15 mg dose.

Of those subjects that entered the double-blind treatment period, approximately half were opioid experienced, and the majority primary pain diagnosis was low back pain 72%. At baseline, the average API was 3.8 and the average WPI was 4.5.

In total, 100 subjects (34%) had their API value imputed in the primary analysis, with proportion relatively similar across treatment arms (HER: 34%; Placebo: 33%)

3.2.2.3 Results

Based on the applicant's prespecified analysis, HER failed to demonstrate a statistically significant difference in API compared to placebo at week 12 (b) (4) that included the null value.

For the secondary endpoint WPI at week 12, the estimated change from baseline for the HER group was -0.35 and 0.20 for placebo. The difference (HER - placebo) was -0.54 with 95% CI (-1.02, -0.07) that excluded the null value. Although the difference between groups on WPI was statistically significant at a nominal 5% significance level, this finding can only be interpreted as being hypothesis generating since the trial failed on the primary endpoint.

Table 13. Analysis of week 12 change in API and WPI (Study 3079)

	HER N = 147	Placebo N = 146
API	(b) (4)	
Adj. mean change from baseline		0.14
HER – Placebo		
(95% CI)		
p-value		
WPI		
Adj. mean change from baseline	-0.35	0.20
HER – Placebo	-0.54	
(95% CI)	(-1.02, -0.07)	

** Source: CSR, Tables 22 (API) and 30 (WPI)*

A post hoc analysis of WPI excluding the 15 mg dose was performed overall and by opioid status. Results from these analyses are provided in Table 14. (b) (4)

Table 14. Analysis of week 12 change in WPI excluding 15 mg dose overall and by opioid status (Study 3079)

	HER	Placebo
Overall	(b) (4)	
Adj. mean change from baseline		0.46
HER – Placebo (95% CI)		
Opioid naïve		
Adj. mean change from baseline		0.21
HER – Placebo (95% CI)		
Opioid experienced		
Adj. mean change from baseline		0.03
HER – Placebo (95% CI)		

3.3 Evaluation of Safety

An evaluation of safety for trial 3103 is included in the FDA clinical review by Dr. Robert Levin of the Division of Anesthesia, Analgesia, and Addiction Products.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

This section summarizes results from the analysis WPI change at week 12 within subgroup levels. Data collected after a subject discontinued treatment was included in the analysis. The subgroups explored are

- Sex (females; males)
- Age (≤ 65 years; > 65 years)
- Race (white; non-white)
- Opioid status (naïve; experienced)
- Stable HER dose (30 mg; 45 mg; 60 mg; 90 mg)

The subgroup analysis was performed on WPI change at week 12 using an ANCOVA model within each subgroup. The model included as covariates baseline WPI, treatment, center and opioid status (except for the analysis by opioid status).

Table 15 summarizes results from the subgroup analyses. Results from different subgroups were reasonably in-line with the estimate from the overall analysis. The greatest difference between levels for the subgroups explored was for opioid status, with the effect being more pronounced for the opioid experienced group (-0.82) than for the opioid naïve group (-0.28). These differences, as with all subgroup comparisons, should be interpreted cautiously for several reasons including multiplicity considerations and the fact that the trials were not designed to detect differences across levels of the subgroups.

Table 15. Subgroup analysis of WPI change at week 12 (Study 3103)

	Adj. mean change	Difference: HER - Placebo (95% CI)	Adj. mean change	Difference: HER - Placebo (95% CI)
Sex	**Males**		**Females**	
HER	0.07		0.20	
Placebo	0.44	-0.37 (-0.89, 0.15)	0.80	-0.59 (-1.18, -0.01)
Age	**≤ 65 years**		**> 65 years**	
HER	0.31		-0.36	
Placebo	0.69	-0.38 (-0.80, 0.04)	0.50	-0.86 (-1.90, 0.17)
Race	**White**		**Non-white**	
HER	0.47		-0.42	
Placebo	0.98	-0.51 (-0.98, -0.04)	0.03	-0.45 (-1.08, 0.18)
Opioid Status	**Naïve**		**Experienced**	
HER	0.02		0.22	
Placebo	0.30	-0.28 (-0.81, 0.26)	1.04	-0.82 (-1.35, -0.28)
Stable dose	**30 mg**		**45 mg**	
HER	0.07		0.10	
Placebo	-0.19	-0.26 (-0.99, 0.46)	1.05	-0.95 (-1.61, -0.30)
	60 mg		**90 mg**	
HER	0.72		0.72	
Placebo	1.14	-0.42 (-1.27, 0.44)	0.92	-0.19 (-1.46, 1.07)

5 SUMMARY AND CONCLUSIONS

5.1 Collective Evidence

The amount of missing data coupled with the marginal effect in those with week 12 data does not lead to robust evidence in favor of Vantrela providing greater relief of low back pain than placebo, as measured by the change from baseline in the weekly average WPI at week 12.

A limitation of the applicant's primary analysis is it attempts to describe a hypothetical rather than real-world drug effect. The primary analysis attempts to estimate the effect assuming all subjects could adhere to study treatment for 12 weeks, contrary to the fact that some subjects could not adhere. In addition to the inference not being consistent with the design of trial 3103 (since were subjects were allowed to discontinue treatment), I question whether the analysis meets the statutory requirement of being able to establish "the drug will have the effect it purports or is represented to have."

The number of subjects with week 12 WPI data was greater in the HER group than placebo (84% vs. 81%). Week 12 data measured after subjects discontinued study treatment were excluded from the applicant's primary analysis. Therefore, the number of subjects with week 12 data that contributed to the analysis was less than the number of subjects with week 12 data. In total, 80% in the HER group and 74% in placebo had week 12 data and did not discontinue study drug.

In a supportive analysis that included data subjects discontinued study treatment Vantrela provided a greater relief of low back pain in placebo at week 12. However, a systematic evaluation into the potential impact of missing data on the difference in week 12 WPI change regardless of treatment adherence revealed the overall results could be susceptible to violations in assumptions about the missing data. There were scenarios investigated that could possibly describe the experience of those with missing data (e.g., equal WPI change in the treatment groups among those with missing data) that caused the results to no longer be statistically significant.

Findings on WPI from the failed study (trial 3079) do not provide independent substantiation of the findings from trial 3103 as they were hypothesis generating.

5.2 Labeling Recommendations

Statistical comments on Section 14 of the applicant's proposed label are included in the Table below

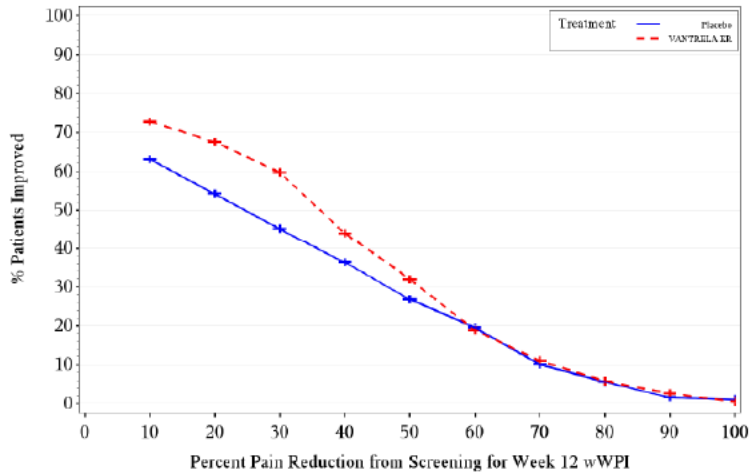
Table 16. Applicant's proposed label and comments

Proposed label	Comment
14 CLINICAL STUDIES The efficacy and safety of VANTRELA ER have been evaluated in a randomized double-blind, placebo-controlled, multi-center clinical trial in opioid-naïve and opioid-experienced patients with moderate to severe chronic low back pain who required continuous opioid treatment for an extended period of time. (b) (4) (b) (4) (b) (4) Opioid-naïve patients initiated VANTRELA ER therapy at the dose 15 mg every 12 hrs. Opioid-experienced patients initiated VANTRELA ER therapy at a dose of hydrocodone that was approximately equivalent to 50% of the dosage of their pre-study opioid medication. Opioid-naïve patients were defined as those taking tramadol or less than 10 mg per day of oxycodone or equivalent during the 14 days before study entry. Opioid-experienced patients were defined as those taking 10 mg or more per day of oxycodone or equivalent but not more than a total of 135 mg/day of oxycodone or equivalent during the 14 days before study entry. (b) (4) A total of 625 patients (mean age = 51.7 [range 19.0-80.0]; 46% male and 54% female) with moderate to severe chronic low back pain were enrolled in the open-label titration period. Of 625 patients enrolled, 371 patients achieved a successful dose and were random (b) (4) VANTRELA ER (191 patients) or placebo (180 patients) during the double-blind treatment period. A total of (b) (4) (b) (4) (b) (4) 277 ((b) (4) patients completed the study. The mean age of patients was 52 years (range 20 to 80 years). The majority (71%) of these patients was white, and the percentages of men and women were similar (49% men and 51% women).	(b) (4) -Should also summarize the number that completed the 12 week treatment period. - The numbers that completed the study and the 12 week treatment period should be summarized by treatment arm.

(b) (4)

The percentage of subjects in each group who demonstrated improvement in their average WPI score at week 12, as compared to Screening is shown in Figure (b) (4). The figure is cumulative, so that patients whose change from screening is, for example, 30%, are also included at every level of improvement below 30%. Patients who did not complete the study were classified as non-responders. The proportion of patients who had up to 30% reduction in their average WPI was higher for the VANTRELA ER treatment group compared with the placebo treatment (b) (4) % vs 45%).

Figure 7: Percentage Improvement in Average Worst Intensity (WPI) Score From Screening to Final Visit at Week 12



(b) (4)

- Remove findings for API

- To make the numbers and plot more interpretable, patients that did not adhere to study drug to week should be classified as non-responders.

-The first sentence of the last paragraph and Figure 7 should clarify that patients that did not adhere to treatment up to week 12 are considered failures.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

BRADLEY W MCEVOY
09/11/2015

FREDA COONER
09/11/2015
I concur.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	NDA 207975/S000 (reference IND 105587)
Drug Name:	hydrocodone Bitartrate Extended-release Tablets (CEP-33237)
Indication(s):	Management of (b) (4) severe pain
Applicant:	Teva Branded Pharmaceutical Products R&D, Inc. (Teva)
Date(s):	Received 9/30/2014; PDUFA due date 10/27/2015
Review Priority:	Standard review
Biometrics Division:	Division of Biometrics VI/Office of Biostatistics (DBVI)
Statistical reviewer:	Feng Zhou, M.S. (DBVI)
Statistics team leader:	Qianyu Dang, Ph.D. (DBVI)
Biometrics division director:	Yi Tsong, Ph.D. (DBVI)
CSS Reviewer:	Katherine Bonson, Ph.D. (CSS)
Project Manager:	Sandra L. Saltz (CSS)
Medical Division:	Division of Anesthetic, Analgesia and Addiction Products (DAAAP)

Keywords: Clinical studies, NDA review, drug abuse

Table of Contents

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

The applicant, Teva Branded Pharmaceutical Products R&D, Inc. (Teva), submitted the results from two abuse potential studies C33237-AP-10032 and C33237/1085 in support of the claim of lack of abuse potential of extended-release hydrocodone product in both intranasal and oral administration routes. These two studies were performed using the methodologies recommended in the 2013 FDA Guidance for Industry: Abuse-Deterrent Opioids – Evaluation and Labeling (http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm334743.pdf).

The design of the two studies was similar; they were randomized, double-blinded, placebo- and active-controlled studies. Study C33237-AP-10032 was conducted to evaluate the abuse potential of manipulated intranasal CEP-33237 compared to intranasal hydrocodone API, placebo, and manipulated intranasal Zohydro™ (commercially available extended-release hydrocodone) through intranasal administration route in healthy, recreational opioid users with a history of intranasal abuse. Study C33237/1085 was conducted to evaluate the abuse potential of the hydrocodone bitartrate extended-release (ER) tablet (CEP-33237) compared to immediate-release (IR) hydrocodone bitartrate and placebo through oral administration rout in healthy, recreational users with a history of opioid abuse.

In intranasal study, a significantly lower mean Emax of Drug Liking VAS and Overall Drug Liking VAS were observed following administration of intranasal CEP-33237 compared to intranasal hydrocodone API, using responses to Drug Liking VAS over 48 hours; the LS mean difference was -6.8 ($p=0.0042$) in Emax of Drug Liking and -8 ($p=0.0044$) in Overall Drug Liking VAS. In contrast, slightly higher Emax scores were observed on both endpoints following administration of intranasal Zohydro™ as compared to intranasal hydrocodone API, although these differences were not statistically. Consistent with this, significantly lower scores on both endpoints were observed following administration of intranasal CEP-33237 compared to intranasal Zohydro™(LS mean difference was -9.9 ($p<0.0001$) in Emax of Drug Liking and -11.7 ($p<0.0001$) in Overall Drug Liking VAS. All 3 intranasal treatments were associated with significantly higher scores than placebo on both endpoints. Significantly higher scores were also observed with intranasal CEP-33237 compared to oral intact CEP-33237, consistent with the similarity of the oral intact tablet to placebo. The majority of subjects were able to completely insufflate the study treatments in all treatment groups (mean and median values >90%).

In oral study, a significantly lower Emax of Drug Liking was observed following administration of the extended-release hydrocodone tablet crushed, as compared with the immediate-release product, based on Emax, using responses to Drug Liking VAS over 72 hours (mean scores of 53.9 and 66.9 vs 85.2; $p<0.001$, respectively). The comparison with placebo showed that mean drug liking VAS Emax for the extended-release tablet intact was not significantly different than placebo (53.9 vs 53.2, $p=0.675$); however the extended-release tablet crushed was significantly different from placebo (66.9 vs 53.2, $p<0.001$). The Emax of Drug Liking was also significantly different following administration of the extended-release tablet intact as compared with the extended-release tablet crushed (53.9 vs 66.9; $p<0.001$).

1 INTRODUCTION

1.1 Overview

1.1.1 Background Information

This consult review responds to a request by the Controlled Substance Staff (CSS) to review the study reports of two human abuse potential studies (oral and intranasal) for CEP-33237 (NDA 207975), an extended-release (ER) hydrocodone single entity drug product that is seeking an abuse deterrent claim. The studies were conducted under IND 105587.

CEP-33237 is an extended-release hydrocodone bitartrate tablets (45 and 60 mg) for the treatment of (b) (4) severe pain. (b) (4) The Sponsor claims this product has tamper deterrence, based on their assertion that their tablet is resistant to rapid release of the drug when the tablet is crushed, and is resistant to dose dumping when co-administered with alcohol. The Sponsor states that, "Although the current abuse patterns for hydrocodone are not predominantly intranasal, CEP-33237 is a single-entity hydrocodone product, and it has the potential to change the pattern of abuse for hydrocodone.

1.1.2 Specific Studies Reviewed

The applicant, Teva, submitted a list of preclinical and clinical study reports related to abuse potential assessment that were conducted and cited in the NDA submission. I will only focus on two human abuse potential studies C33237/1085 (hereafter refer to as C-1085) and C33237-AP-10032 (hereafter refer to as C-10032) in my review after discussed with the CSS reviewer (Table 1). In this review, intranasal, immediate-release, extended-release, hydrocodone, CEP-33237 Active Pharmaceutical Ingredient refers to as IN, IR, EX, Hydro, CEP, and API in tables and figures, respectively.

Table 1: List of Studies Included in this Review

Study ID (Date of submission)	Location	Design	Primary Endpoints	Treatments	Number of Subjects
C-10032 (12/23/2014)	1 site in the USA	R, DB, AC, PC, MD, 5-arms crossover to evaluate the abuse potential of intranasally administered drug	Emax for Drug Liking & Overall Drug Liking VAS	IN CEP-33237 45 mg IN hydrocodone API 45 mg Oral CEP-33237 45 mg Placebo IN Zohydro™ 45 mg	45 randomized and 34 subjects completed all treatment periods
C-1085 Phase I (9/30/2014)	1 site in the USA	R, DB, AC, PC, MD, 4-arms crossover to evaluate the abuse potential of orally administered drug	Emax for Drug Liking VAS	Oral CEP-33237 45 mg (crushed) Oral CEP-33237 45 mg (intact) Hydrocodone IR 45 mg Placebo	49 randomized and 35 subjects completed all treatment periods

Abbreviations: DB = double blind; PC = placebo-controlled; AC = active-controlled; R = randomized; MD=multi-dose

1.2 Data Sources

All data was supplied by the applicant to the CDER electronic data room in SAS transport format. The data and final study report for the electronic submission were archived under the network path location \\...\207975.enx. The information needed for this review was contained in submission modules 5.3.4 modules and datasets.

2 STATISTICAL EVALUATION

2.1 Data and Analysis Quality

Data for study C-1085 was submitted on 9/30/2014 (NDA 207975/S0000); Data for study C-10032 was submitted on 12/23/2014 (NDA 207975/S0001). In general, the data and analysis quality are acceptable.

2.2 Human Abuse Potential Stud C-10032

2.2.1 Overview

C-10032 was a single-dose, double-blind, randomized crossover study to assess the intranasal (IN) pharmacokinetics, abuse potential and safety of CEP-33237 compared to IN hydrocodone API in healthy, nondependent, recreational opioid users with a history of intranasal abuse.

2.2.1.1 Objectives of the Study

The primary objective of the study was to assess the relative abuse potential of manipulated IN CEP-33237 as compared to that of IN hydrocodone Active Pharmaceutical Ingredient (API) (a surrogate for the immediate-release product), based on the peak score (Emax) of the Drug Liking visual analog scale (VAS) and Overall Drug Liking VAS Emax.

Secondary Objectives are mainly to assess the relative abuse potential of manipulated IN CEP-33237 as compared to IN hydrocodone API, oral intact CEP-33237, or manipulated IN Zohydro™ (commercially available ER hydrocodone), as assessed by all primary and secondary pharmacodynamics variables.

2.2.1.2 Study Design

This is a single-dose, randomized, double-blind, quadruple-dummy, active and placebo controlled crossover study designed to assess the abuse potential of manipulated IN CEP-33237 in healthy nondependent, recreational opioid users.

The study consisted of 3 phases:
Screening Phase (A): at visit 1, subjects where evaluated to determine if they met criteria for inclusion and exclusion.
Qualification Phase (B): at visit 2, eligible subjects randomized to placebo and IN hydrocodone API 45 mg to ensure tolerability and appropriate reporting of positive subjective effects.
Treatment Phase (C): was the randomized, double-blind, quadruple-dummy, placebo-controlled, 5-period crossover treatment portion of the study. There were 5 visits during this phase (Visit 3-7). Eligible subjects were randomly assigned to 1 of 10 treatment sequences according to two 5x5 Williams squares, defined as follows (Table 2):

ABECD, BCADE, CDBEA, DECAB, EADBC,
DCEBA, EDACB, AEBDC, BACED, CBDAE

The Follow-up Phase (D): subjects visited approximately 48-72 hours after the last discharge from the study center/final assessment or early withdrawal.

The eligible subjects were male or female recreational opioid users, 18 to 55 years of age, who were nondependent recreational opioid users with intranasal experience. Approximately 45 subjects were planned for randomization to the treatment phase to ensure that at least 30 subjects completed all Treatment visits.

In the treatment phase, subjects received each of the treatments once (Table 2). There was a minimum 7 day washout period between each administration of study drug.

Table 2: Summary of Treatments Administered

Treatment Identifier	Intranasal Treatments (Each active treatment consisted of a 45 mg dose of hydrocodone from the specified product administered in a total weight of ~575 mg of material)			Oral Treatments
	Container 1	Container 2	Container 3	
A (intranasal CEP-33237)	~90 mg of manipulated 45 mg CEP-33237 tablet	~158 mg of manipulated 45 mg CEP-33237 tablet	~327 mg of manipulated 45 mg CEP-33237 tablet	1 intact CEP-33237 placebo tablet
B (intranasal hydrocodone API)	~45 mg hydrocodone API plus ~45 mg lactose	~158 mg crushed sugar spheres	~327 mg lactose	1 intact CEP-33237 placebo tablet
C (oral CEP-33237)	~90 mg crushed sugar spheres	~158 mg lactose	~327 mg crushed sugar spheres	1 intact 45 mg CEP-33237 tablet
D (placebo)	~90 mg manipulated CEP-33237 placebo tablet	~158 mg manipulated CEP-33237 placebo tablet	~327 mg manipulated CEP-33237 placebo tablet	1 intact CEP-33237 placebo tablet
E (intranasal hydrocodone extended-release capsules, (Zohydro™)	~90 mg of manipulated contents of 2 hydrocodone extended-release capsules (one 15 mg and one 30 mg capsule)	~158 mg of manipulated contents of 2 hydrocodone extended-release capsules (one 15 mg and one 30 mg capsule)	~327 mg lactose	1 intact CEP-33237 placebo tablet

[Source: Table 2 of study report report-body.pdf, page 32]
Note: shading is intended to show where particle size/powder was most similar (eg, sugar spheres placebo and manipulated ZohydroTM, manipulated CEP-33237 and CEP-33237 placebo tablets, and HYDR API and lactose).

2.2.1.3 Abuse Potential Measures

Pharmacodynamic (PD), pharmacokinetic (PK), and safety assessments were conducted up to 24 hours post-dose. All subjects were discharged from each visit after completion of the final (i.e., 24 hours) post-dose procedures. Each study drug administration was separated by a minimum of 72 hours.

Primary Endpoint consisted of the Emax for Drug Liking VAS [question 1 of the Drug Liking and Effects Questionnaire (DLEQ)] and Emax of the Overall Drug Liking VAS.

Secondary measures were included to evaluate other subjective effects including balance of effects (Take Drug Again VAS); positive effects (High VAS and Good Effects VAS); negative effects (Bad Effects VAS, ARCI Lysergic Acid Diethylamide [LSD], and Subject-rated Assessment of Intranasal Irritation [SRAII]); sedative effects (ARCI Pentobarbital and Chlorpromazine Alcohol Group [PCAG] and Alertness/Drowsiness VAS); and other drug effects (Any Drug Effects VAS). Observer-related Assessment of Intranasal Irritation (ORAII) using endoscopy was also conducted as was the objective measure of pupillometry.

The DLEQ and pupil diameter measurements were completed at 0.25, 0.5, 0.75, 1, 1.25, 1.5, 2, 2.5, 3, 4, 6, 7, 8, 9, 10, 12, 24, 36, and 48 hours post-dose; pupil diameter measurements were also collected at pre-dose. The Overall Drug Liking VAS, the Take Drug Again VAS, and the PVAQ assessment were completed at 12 and 24 hours post-dose during in each period of the treatment phase (See Table 15 in Appendix for the detail).

2.2.1.4 Analysis Population and Sample Size

For subjects randomized in the treatment phase, the mean age was 27.8 years (range, 19.0 to 52.0 years); 73% of subjects were men and 87% were white. Mean BMI was 23.6 kg/m2 (range, 18.4 to 30.4 kg/m2). (See Table 14 in Appendix for the detail.) The primary analysis was based on the pharmacodynamics population (hereafter referred to as the PD analysis set) that included all subjects who received all 5 treatments in the treatment phase. No imputation was performed for any missing measurements. A total of 45 subjects were randomized to the Treatment Phase and 34 subjects completed all 5 treatment periods (See Table 16 in Appendix for the detail).

The sponsor claimed a sample size of 30 completed subjects will have at least 90% power to detect a difference of 15 points on VAS scale (0-100 points) between a pair of treatment, assuming an estimated within-subjects standard deviation of 17.5 using a paired t-test with a 2-sided significance level of 5%.

2.2.1.5 Statistical Methodologies used in the Sponsor's Analyses

The sponsor's primary endpoints and other continuous and ordinal categorical pharmacodynamics parameters were analyzed using a mixed-effects analysis of variance (ANOVA) model for a crossover study. The model included: treatment, period, and randomized treatment sequence as fixed effects, baseline (pre-dose) measurement as covariate where applicable, and subject nested within sequence as a random effect. The first order carryover was included in the model as fixed effect and was to be dropped if not statistically significant at the 25% significance level. The residuals from the mixed-effect model were investigated for normality using the Shapiro-Wilk W-test. Parameters were analyzed as having a normal distribution if the probability value is ≥ 0.05. Parameters that did not meet this criterion were analyzed non-parametrically. Overall treatment effect was assessed using Friedman's test; pairwise treatment comparisons were assessed using the Wilcoxon sign-rank test on the within-subject differences. The comparison of IN CEP-33237 vs IN hydrocodone API was performed only if statistical significant difference was observed between IN hydrocodone API vs placebo for both primary endpoints. In order to claim less abuse liability for IN CEP-33237 compared to IN hydrocodone API, statistical significant treatment difference had to be observed for both

primary endpoints in favor of IN CEP-33237. For each of the pairwise comparisons, the null hypothesis is: there is no treatment effect difference between the tested pair, and the alternative hypothesis is: there is a treatment effect difference between the tested pair. A 5% Type-I error rate with a p-value less than 0.05 will be considered as statistically significance for all individual hypothesis tests. All statistical tests will be performed using two-tailed significance criteria. Secondary endpoints were analyzed in a similar manner as the primary endpoint. No adjustments will be made for the preplanned multiple comparisons and endpoints.

The percent reduction in Emax was calculated for Drug Liking VAS as follow:

$$\%reduction = \begin{cases} \frac{C-T}{C-50} \times \left(1 - \frac{P-50}{50}\right) \times 100\%, & if\ P > 55; \\ \frac{C-T}{C-50} \times 100\%, & if\ P \le 55. \end{cases}$$

2.2.1.6 Changes in the Conduct of the Study

The Statistical Analysis Plan was finalized on August 25, 2014. There were some changes in the SAP prior to database lock and un-blinding. These changes did not effect on the primary analyses.

2.2.1.7 Sponsor's Summary and Conclusions

The sponsor summarized that:

- *There were statistically significant differences between placebo and IN hydrocodone API for the primary endpoints of Drug Liking VAS Emax and Overall Drug Liking VAS Emax, thereby confirming study validity.*
- *IN CEP-33237 was associated with significantly lower effects compared to IN hydrocodone API and IN Zohydro™ on the primary endpoints, as well as secondary balance of effects, positive, sedative and other effects endpoints. IN CEP-33237 showed greater peak "bad effects" compared to IN hydrocodone API but not IN Zohydro™. In contrast, IN Zohydro™ was associated with similar or greater effects on the primary and secondary pharmacodynamic endpoints compared to IN hydrocodone API. IN CEP-33237 also showed a slower onset of effects compared to IN hydrocodone API and IN Zohydro™.*
- *All active IN treatments were associated with significantly greater effects in comparison to placebo on the primary and most secondary endpoints, while oral intact CEP-33237 showed subjective effects similar to placebo. Consistent with the similarity between oral intact CEP-33237 and placebo, statistical comparisons of IN CEP-33237 and oral intact CEP-33237 showed significant differences on most endpoints.*

The sponsor concluded that:

Consistent with the observed differences in pharmacokinetics across treatments, IN CEP 33237 demonstrated significantly lower subjective effects compared to IN hydrocodone API and Zohydro™, with a markedly different timecourse profile of slower onset and rate of rise, lower peak and shorter duration of effects. While not evaluated statistically, oral intact CEP-33237 showed a similar pharmacodynamic profile to placebo. Overall, the pharmacodynamic and pharmacokinetic results demonstrate that CEP 33237 may have lower abuse potential compared to non-abuse deterrent opioid products, including Zohydro™.

2.2.2 Reviewer's Assessment

This reviewer focused on the primary endpoints and selected endpoints: Emax of nasal effect VAS and pupil diameter. All analyses were based on completer population.

2.2.2.1 Primary Endpoints – Drug Liking Emax and Overall Drug Liking Emax

Descriptive statistics show that the mean and median Drug Liking Emax scores and Overall Drug Liking Emax for the subjects treated with IN hydrocodone API and IN Zohydro™ were relatively high (approximately 79 or higher) (Table 3). In contrast, mean and median of two endpoints for the subjects treated with IN CEP-33237 was 73 or lower. Mean and median Emax scores for the subjects treated with oral intact CEP-33237 were similar to placebo for both endpoints. The similar tables for High VAS, ARCI MBG, Good Effect VAS, Bad Effect VAS, Overall Drug Liking VAS, and Take Drug Again VAS can be found in Table 17 in the Appendix.

Figure 1 shows that the mean Drug Liking VAS scores for the subjects treated with IN hydrocodone API and IN Zohydro™ increased rapidly, with a steep rate of rise and an onset of effects beginning at approximately 0.5 hours post-dose. Both treatments were associated with relatively high mean scores (>65) for 3.5 hours (from 0.5 hours until at least 4 hours post-dose). However, IN hydrocodone API showed a defined peak at 1.5 hours post-dose (mean of 76.2) while IN Zohydro™ showed a more prolonged peak, with scores slightly higher than IN hydrocodone API between 0.75 and 2.0 hours post-dose. In contrast, IN CEP-33237 treated subjects was associated with a slower rise in Drug Liking VAS scores to a lower peak effect (up to a mean of 67.9 at 2 hours post-dose). Mean scores were above 65 for a shorter period of time, later in the time-course profile (from 2 to 3 hours post-dose) as compared to IN API and IN Zohydro™. Drug Liking VAS scores over time were comparable following administration of oral intact CEP-33237 and placebo. Both showed little increase above the neutral point of the scale. Drug Liking VAS scores did not show much change after 12 hours post-dose (Figure 12 in Appendix).

In the heat map (Figure 2), the density of the color blue indicates the degree of the disliking and the density of the color red indicates the degree of the liking. This figure shows that most of subjects treated with IN CEP-33237 doses were neutral and slightly above neutral except there were few subjects liked. Notice that, 11 out of 34 subjects had high placebo response (Emax of Drug Liking >60), this may be one of the reasons that there was no significant difference in Emax between oral intact CEP-33237 and placebos; similarly, 8 out of 34 subjects had Overall Drug Liking VAS scores > 60 (Figure 3).

Inferential Statistics show that significantly lower Emax scores for Drug Liking VAS and Overall Drug Liking VAS were observed following administration of IN CEP-33237 compared to IN hydrocodone API, the mean difference was -6.83 (p=0.0042) and -8.02 (p=0.0044) for Emax scores for Drug Liking VAS and Overall Drug Liking VAS, respectively (Table 4). In contrast, slightly higher Emax scores were observed on both endpoints following administration of IN Zohydro™ as compared to IN hydrocodone API, although these differences were not statistically significant. Consistent with this, significantly lower scores on both endpoints were

observed following administration of IN CEP-33237 compared to IN Zohydro™. All 3 IN treatments were associated with significantly higher scores than placebo on both endpoints. Significantly higher scores were also observed with IN CEP-33237 compared to oral intact CEP-33237.

Table 3: Summary Statistics for Emax of Drug Liking VAS (N=34)

Treatments	N	Mean	StdErr	Min	Q1	Med	Q3	Max
Emax - Drug Liking VAS (DLEQ-Q1)								
IN CEP-33237 45 mg	34	72.8	2.35	50	63	72.5	82	100
IN Hydro API 45 mg	34	80.2	2.16	57	73	79.0	88	100
Oral-CEP-33237 45 mg	34	57.3	1.88	50	50	52.0	61	93
PLACEBO	34	58.7	1.94	50	50	52.0	66	90
IN ZOHYDRO 45 mg	34	83.2	2.04	59	74	84.0	94	100
Overall Drug Liking VAS								
IN CEP-33237 45 mg	34	68.5	3.31	34	56	71.5	81	100
IN Hydro API 45 mg	34	77.1	2.52	42	69	77.0	85	100
Oral-CEP-33237 45 mg	34	57.8	2.69	31	50	50.0	58	99
PLACEBO	34	57.7	2.39	38	50	50.0	60	98
IN ZOHYDRO 45 mg	34	79.8	2.72	34	71	80.0	94	100

Table 4: The Treatment Comparison in LS Means of Emax of Drug Liking (VAS) (N=34)

Treatment Comparison	LS Mean Diff.	Std Error	P-value	LL-95%CI	UL-95%CI
Emax - Drug Liking VAS (DLEQ-Q1)					
IN CEP 45 mg – IN Hydro API 45 mg	-6.83	2.34	0.0042	-11.47	-2.19
IN CEP 45 mg – IN ZOHYDRO 45 mg	-9.92	2.32	<.0001	-14.52	-5.32
IN ZOHYDRO 45 mg – IN Hydro API 45 mg	3.09	2.34	0.1894	-1.54	7.73
IN CEP 45 mg - Placebo	12.51	2.33	<.0001	7.89	17.12
IN Hydro API 45 mg - Placebo	19.34	2.31	<.0001	14.76	23.92
IN ZOHYDRO 45 mg - Placebo	22.43	2.33	<.0001	17.82	27.04
Oral-CEP 45 mg - Placebo	-2.88	2.33	0.2199	-7.50	1.74
IN CEP 45 mg – Oral-CEP 45 mg	15.39	2.34	<.0001	10.75	20.02
Overall Drug Liking VAS					
IN CEP 45 mg – IN Hydro API 45 mg	-8.02	2.77	0.0044	-13.50	-2.55
IN CEP 45 mg – IN ZOHYDRO 45 mg	-11.67	2.74	<.0001	-17.10	-6.25
IN ZOHYDRO 45 mg – IN Hydro API 45 mg	3.65	2.76	0.1889	-1.82	9.12
IN CEP 45 mg - Placebo	8.99	2.75	0.0014	3.55	14.43
IN Hydro API 45 mg - Placebo	17.02	2.73	<.0001	11.62	22.42
IN ZOHYDRO 45 mg - Placebo	20.67	2.75	<.0001	15.22	26.11
Oral-CEP 45 mg - Placebo	-0.53	2.75	0.8485	-5.98	4.92
IN CEP 45 mg – Oral-CEP 45 mg	9.52	2.76	0.0008	4.05	14.99

Figure 1: Mean Drug Liking VAS Over Time (0-12 hours) by Treatment (N=34)

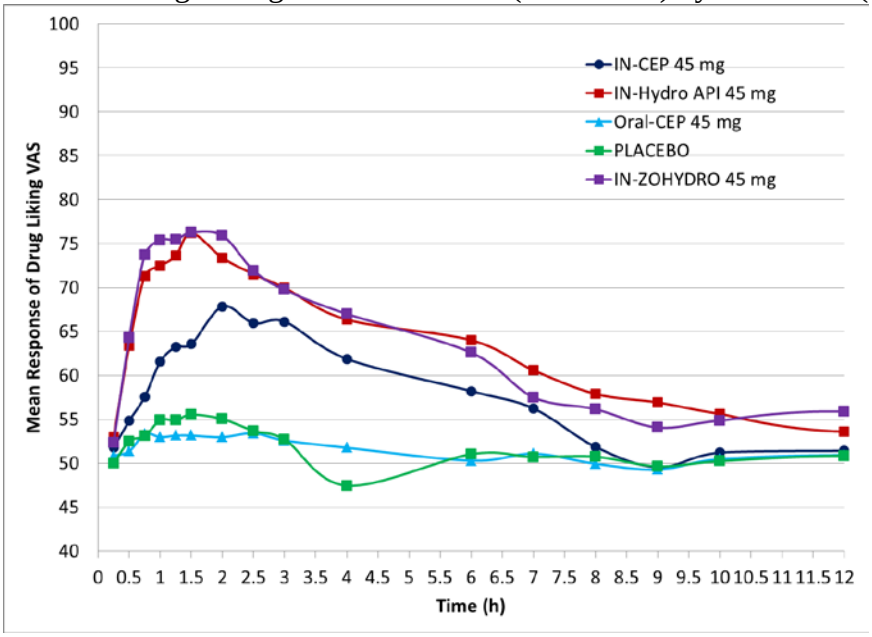


Figure 2: Heat Map for Emax of Drug Liking VAS by Treatment (N=34)

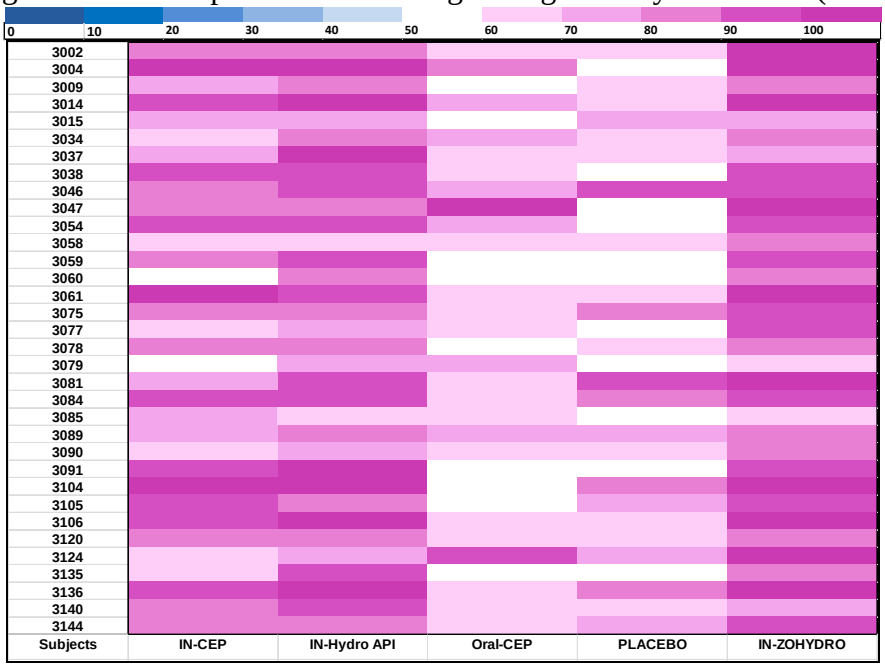
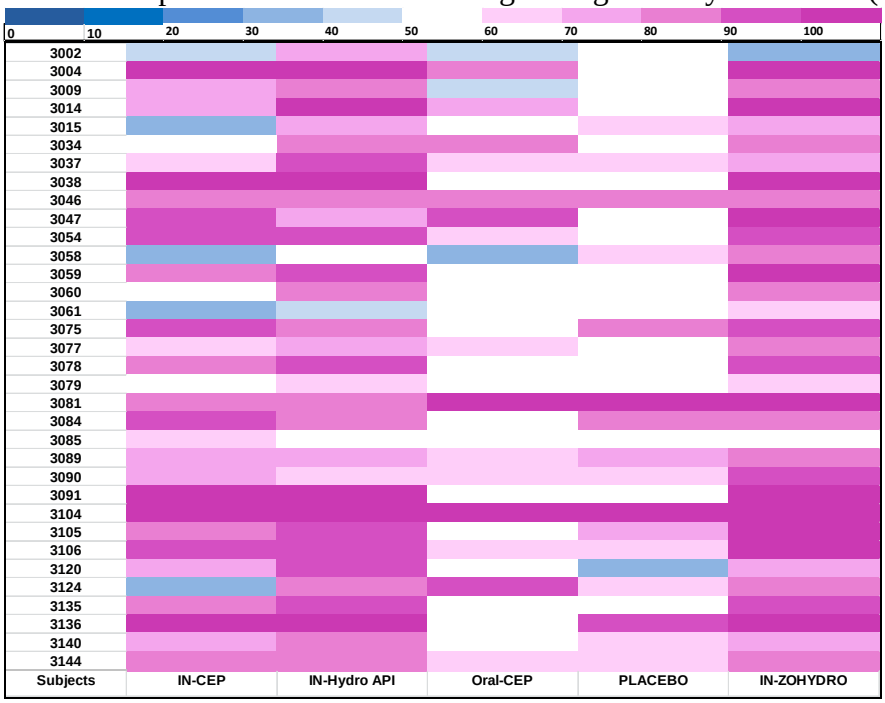


Figure 3: Heat Map for Emax of Overall Drug Liking VAS by Treatment (N=34)



Statistical issues and post hoc analyses:

1. A significant first order carryover effect was observed for both endpoints (p=0.0001). The interaction of treatment by period was not statistical significant. The evaluation of data from the first period indicates that, for most treatments, the first period data were comparable to data for the completers analysis set (periods 1 - 4 combined). The statistical analysis results were based on all period's data.

2. The assumption of normal distribution of the data was examined using Shapiro-Wilk Normality test. Since the probability value was ≤ 0.05, the pairwise treatment comparisons were assessed using the Wilcoxon signed-rank test on the within-subject differences (Table 5). The results are similar to the results from mixed effects model.

3. Responder analysis results also support the primary analysis results. Given median of Emax of Drug Liking VAS greater than 78 for IN API (79) and IN Zohydro™ (84), approximately 17.6% and 23.5% of subjects had at least 50% reduction for IN CEP-33237 relative to IN API and IN Zohydro™, respectively (Figure 4). For Overall Drug Liking VAS approximately 29.4% of subjects had at least 50% reduction for IN CEP-33237 relative to IN API and IN Zohydro™ (Figure 5).

Table 5: The Treatment Comparison in Median of Emax of Drug Liking (VAS) (N=34)

Overall Treatment Effect	Median Differ.	Inner Quarter Range		P-value
Emax - Drug Liking VAS (DLEQ-Q1)				
IN Hydro API 45 mg - Placebo	21.0	8	32	0.00000
IN ZOHYDRO 45 mg - Placebo	21.5	14	36	0.00000
IN CEP 45 mg - Placebo	12.5	2	27	0.00000
Oral-CEP 45 mg - Placebo	0.0	-12	5	0.56316
IN CEP 45 mg – IN Hydro API 45 mg	-6.0	-14	-1	0.00005
IN CEP 45 mg – IN ZOHYDRO 45 mg	-9.5	-19	-1	0.00000
IN ZOHYDRO 45 mg – IN Hydro API 45 mg	1.5	-2	8	0.07155
IN CEP 45 mg – Oral-CEP 45 mg	17.5	7	27	0.00001
Overall Drug Liking VAS				
IN Hydro API 45 mg - Placebo	17.5	6	32	0.00000
IN ZOHYDRO 45 mg - Placebo	21.0	9	32	0.00000
IN CEP 45 mg - Placebo	9.5	0	26	0.00331
Oral-CEP 45 mg - Placebo	0.0	-5	4	0.96688
IN CEP 45 mg – IN Hydro API 45 mg	-5.0	-15	0	0.00038
IN CEP 45 mg – IN ZOHYDRO 45 mg	-6.5	-20	0	0.00000
IN ZOHYDRO 45 mg – IN Hydro API 45 mg	0.0	-3	12	0.31134
IN CEP 45 mg – Oral-CEP 45 mg	7.5	0	26	0.00464

Figure 4: Proportion of Responders for Emax of Drug Liking VAS, Study C-10032 (N=34)

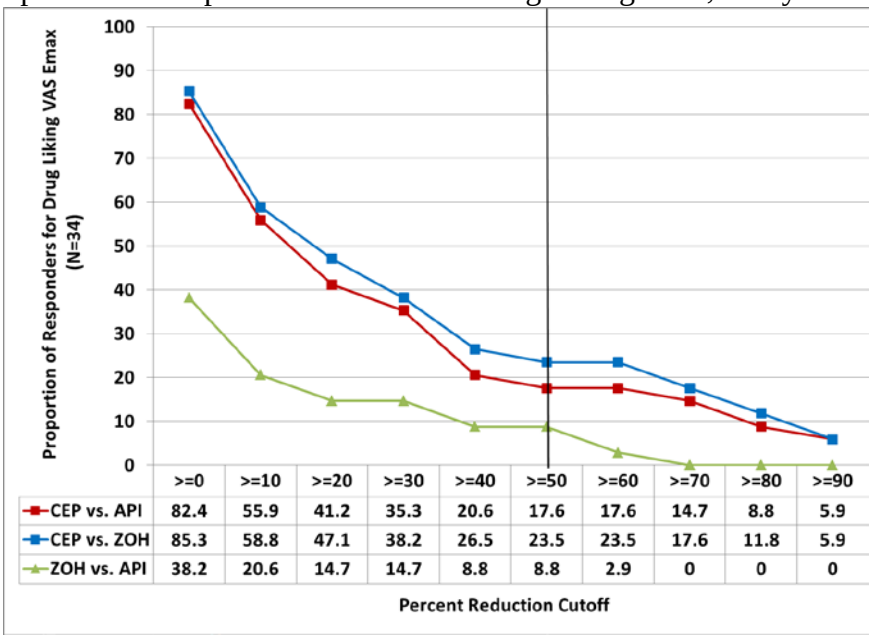
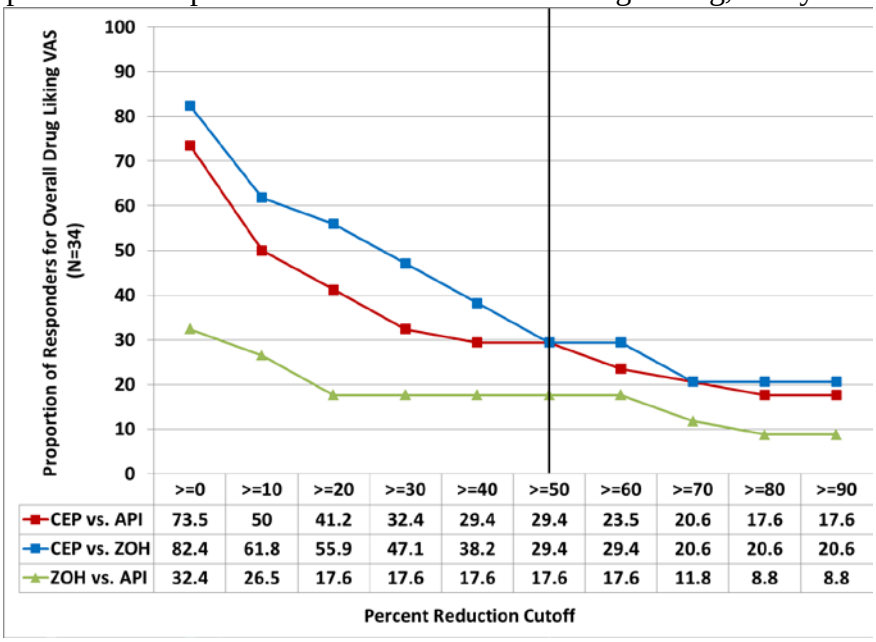


Figure 5: Proportion of Responders for Emax of Overall Drug Liking, Study C-10032 (N=34)



	>=0	>=10	>=20	>=30	>=40	>=50	>=60	>=70	>=80	>=90
CEP vs. API	73.5	50	41.2	32.4	29.4	29.4	23.5	20.6	17.6	17.6
CEP vs. ZOH	82.4	61.8	55.9	47.1	38.2	29.4	29.4	20.6	20.6	20.6
ZOH vs. API	32.4	26.5	17.6	17.6	17.6	17.6	17.6	11.8	8.8	8.8

2.2.2.2 Measures of Nasal Effects

Measures of nasal effects included Ease of Snorting VAS (bipolar scale; "Snorting this drug was" where responses range from 0=Very easy to 100=Very difficult) and Emax and AUEC0-8h of SRAII scales (Burning, Need to Blow Nose, Runny Nose/Nasal Discharge, Facial Pain/Pressure and Nasal Congestion; rated on a 6-point scale from 0=Not observed/No problem, 1=Very Mild Problem; 2=Mild/Slight Problem; 3=Moderate Problem; 4=Severe Problem; to 5=Very Severe Problem/"As Bad as Can Be"). For Ease of Snorting VAS, there was only one statistical significant difference in the pairwise comparisons, between oral intact CEP-33237 and IN CEP-33237 (mean difference 12.6; p=0.017, Table 6).

For SRAII, IN CEP-33237 showed significantly lower effects compared to IN hydrocodone API and IN Zohydro™ on SRAII-Burning Emax and SRAII-Runny Nose/Nasal Discharge Emax. IN CEP-33237 was different from oral intact CEP-33237 on SRAII-Facial Pain/Pressure Emax, Need to Blow Nose Emax, and Nasal Congestion Emax. Compared to placebo, hydrocodone API and Zohydro™ showed greater effects on SRAII-Burning Emax and SRAII-Runny Nose/Nasal Discharge Emax. Other endpoints supported the results of the primary endpoints. (Table 17 in appendix)

Table 6: Selected Descriptive Statistics of Secondary Nasal Effects Endpoints (N=34)

Parameter	Statistics	IN CEP (N=34)	IN Hydro API (N=34)	Oral-CEP (N=34)	Placebo (N=34)	IN ZOHYDRO (N=34)
Ease of Snorting VAS, Emax						
Score	Mean (SE)	42.2 (4.6)	40.5 (4.3)	29.2 (3.8)	32.0 (4.1)	36.0 (4.7)
	Median (Range)	41.0 (0, 91)	43.0 (0, 86)	23.0 (0, 81)	30.0 (0, 93)	29.5 (0, 100)
	LS Mean Diff	--	1.45 (p=0.781)	12.6 (p=0.017)	9.9 (p=0.059)	6.0 (p=0.252)
SRAII – Burning, Emax						
Score	Mean (SE)	0.9 (0.17)	1.3 (0.19)	0.5 (0.15)	0.6 (0.15)	1.6 (0.21)

	Median (Range)	1 (0, 4)	1 (0, 4)	0 (9, 3)	0 (0, 3)	2 (0, 4)
	LS Mean Diff	--	-0.41 (p=0.037)	0.38 (p=0.053)	1.1 (p<0.001)	-0.71 (p<0.001)
SRAII – Need to blow nose, Emax						
Score	Mean (SE)	1.9 (0.21)	1.9 (0.19)	1.4 (0.21)	1.6 (0.21)	2.0 (0.21)
	Median (Range)	2 (0, 4)	2 (0, 4)	1 (0, 4)	1.5 (0, 4)	2 (0, 5)
	LS Mean Diff	--	-0.06 (p=0.790)	0.44 (p=0.048)	0.24 (p=0.289)	-0.12 (p=0.595)
SRAII – Runny nose/nasal discharge, Emax						
Score	Mean (SE)	1.1 (0.19)	1.7 (0.19)	1.3 (0.21)	1.2 (0.17)	1.8 (0.19)
	Median (Range)	1 (0, 4)	2 (0, 4)	1 (0, 5)	1 (0, 3)	2 (0, 4)
	LS Mean Diff	--	-0.65 (p=0.002)	-0.24 (p=0.251)	-0.12 (p=0.568)	-0.74 (p<0.001)
SRAII – Facial pain/pressure, Emax						
Score	Mean (SE)	1.1 (0.2)	1 (0.17)	0.5 (0.16)	0.8 (0.19)	1.2 (0.21)
	Median (Range)	1 (0, 4)	1 (0, 3)	0, (0, 4)	0 (0, 4)	1 (0, 4)
	LS Mean Diff	--	0.06 (p=0.763)	0.56 (p=0.005)	0.26 (p=0.176)	-0.12 (p=0.547)
SRAII – Nasal congestion, Emax						
Score	Mean (SE)	1.8 (0.22)	1.5 (0.19)	1.3) (0.21)	1.9 (0.20)	1.7 (0.22)
	Median (Range)	2 (0, 5)	1 (0, 3)	1 (0, 4)	1.5 (0, 4)	1.5 (0, 4)
	LS Mean Diff	--	0.32 (p=0.123)	0.56 (p=0.008)	-0.09 (p=0.673)	0.15 (p=0.482)
Pupil Diameter - Emin						
Score	Mean (SD)	3.4 (0.59)	3.3 (0.65)	4.0 (0.78)	5.5 (0.75)	3.0 (0.49)
	Median (Range)	3.3	3.3	4.0	5.6	3.1
	LS Mean Diff	--	0.05 (p=0.677)	-0.71 (p<0.001)	-2.18 (p<0.001)	0.30 (p=0.006)

2.2.2.3 Pupil Diameter Endpoint

Pupil diameter Emin was significantly greater following administration of IN CEP-33237 (indicating less pupillary constriction) compared to IN Zohydro™, but was not significantly different from IN hydrocodone API (Table 7). The oral intact CEP-33237 was associated with a much slower onset compared to the other active treatments and persisted for a longer duration of time. Placebo did not affect pupil diameter (Figure 6). These findings confirmed the analysis results based on the Drug Liking VAS and Overall Drug Liking VAS. Because there is large number of subjects who had placebo response, it resulted in no significant difference between oral intact CEP-33237 and placebo on Drug Liking VAS and Overall Drug Liking VAS.

Figure 6: The Mean Pupil Diameter Over 12-hours by Treatment, Study C-10032 (N=34)

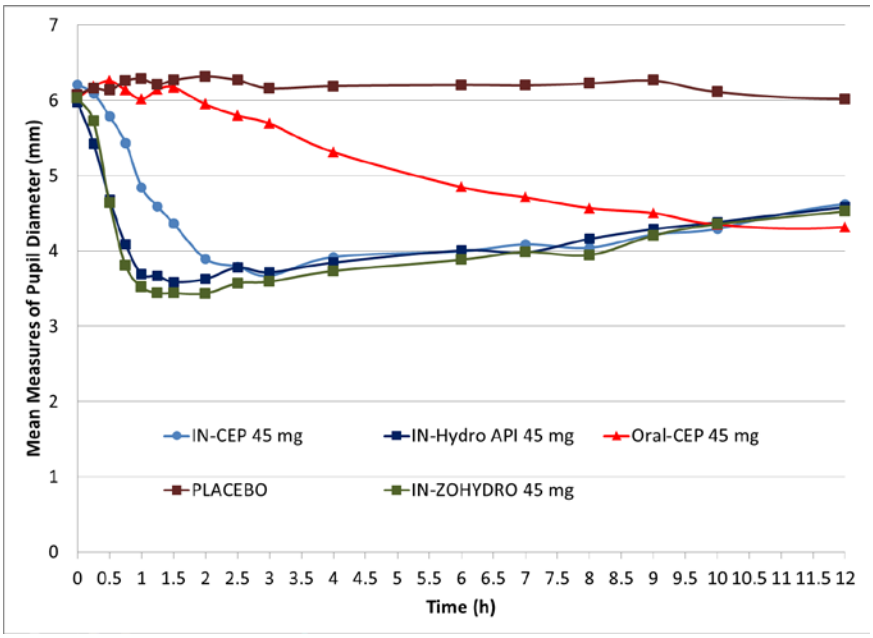


Table 7: Summary Statistics for Emin of Measured Pupillometry (N=34)

Treatment Comparison	LS Mean Diff.	SE	P-Value	Lower of 95% CI	Upper of 95% CI
IN CEP 45 mg – IN Hydro API 45 mg	0.1033	0.1094	0.3466	-0.1131	0.3198
IN CEP 45 mg – IN ZOHYDRO 45 mg	0.3605	0.1084	0.0012	0.1459	0.5750
IN ZOHYDRO 45 mg – IN Hydro API 45 mg	-0.2572	0.1092	0.0201	-0.4734	-0.04097
IN CEP 45 mg - Placebo	-2.1155	0.1087	<.0001	-2.3307	-1.9004
IN Hydro API 45 mg - Placebo	-2.2189	0.1079	<.0001	-2.4325	-2.0053
IN ZOHYDRO 45 mg - Placebo	-2.4760	0.1087	<.0001	-2.6912	-2.2608
Oral-CEP 45 mg - Placebo	-1.4590	0.1089	<.0001	-1.6746	-1.2434
IN CEP 45 mg – Oral-CEP 45 mg	-0.6565	0.1093	<.0001	-0.8728	-0.4402

2.2.2.4 Percentage of Dose Insufflated Endpoint

With the exception of one subject (# 12613056 received all treatments but insufflated only 29% of the IN hydrocodone API dose), the majority of subjects were able to completely insufflate the study treatments in all treatment groups (mean and median values >90%, Table 8).

Table 8: Descriptive Statistics of Percentage of Dose Insufflated

Statistic	Placebo (N=38)	45 mg IN API (N=39)	45 mg IN Zohydro™ (N=39)	45 mg IN CEP-33237 (N=41)	45 mg OR CEP-33237 (N=38)
Mean (SD)	97.7 (1.39)	96.9 (11.20)	97.6 (3.26)	96.8 (1.86)	99.7 (2.06)
Median	97.9	99.0	96.3	96.4	99.4
Min, Max	94.6, 100.7	29.2, 100.9	93.9, 105.6	92.6, 101.1	98.3, 111.8

Source [page 70 in report-body.pdf]

2.2.2.5 Conclusion

The reviewer's statistical analysis showed that the median (or mean) responses to IN CEP-33237 were significantly lower than those to IN API and IN Zohydro™ for Drug Liking VAS and Overall Drug Liking VAS. All 3 intranasal treatments were associated with significantly higher scores than placebo on both endpoints. The heat maps for Drug Liking VAS showed that overall the time course response profiles for individual subjects to IN CEP-33237 were different from those to IN API and IN Zohydro™. Given median of Emax of Drug Liking VAS greater than 78 for IN API (79) and IN Zohydro™ (84), approximately 17.6% and 23.5% of subjects had at least 50% reduction for IN CEP-33237 relative to IN API and IN Zohydro™, respectively.

Even though there were still some subjects who strongly liked IN CEP-33237 administered intranasally in study C-10032, it is clear that the IN CEP-33237 formulation may have the advantage of making some subjects dislike or less like the drug through nasal route and it may have abuse-deterrent effect.

Statistically significant decreases in pupil diameter Emin were observed following all active intranasal treatments in comparison to placebo. Pupil diameter Emin was significantly greater following administration of IN CEP-33237 compared to IN Zohydro™, but was not significantly different from IN hydrocodone API. The oral intact CEP-33237 was associated with a much slower onset compared to the other active treatments and persisted for a longer duration of time. Placebo did not affect pupil diameter, however there is large number of subjects who had placebo response on VAS scales, it resulted in no significant difference between oral intact CEP-33237 and placebo on Drug Liking VAS and Overall Drug Liking VAS.

2.3 Human Abuse Potential Stud C-1085

2.3.1 Overview

C-1085 was a randomized, double-blind, placebo-controlled, crossover, phase I study to assess the abuse potential of the hydrocodone bitartrate ER tablet (here in referred to as CEP-33237) compared to IR hydrocodone bitartrate (herein referred to as IR hydrocodone) and placebo in healthy, nondependent, and recreational opioid users.

2.3.1.1 Objectives of the Study

The primary objectives were to assess the relative abuse potential of the CEP-33237 (crushed or intact) as compared with that of IR hydrocodone based on the Emax of the Drug Liking VAS.

The secondary objectives were to assess the relative abuse potential of the CEP-33237 (crushed or intact) as compared with that of IR hydrocodone as assessed by all secondary pharmacodynamic variables, and to assess the relative abuse potential of the CEP-33237 (crushed) as compared with that of the CEP-33237 (intact) as assessed by the primary and secondary pharmacodynamic variables.

Reviewer's comments: CEP-33237 is an ER product. With the same dose of hydrocodone, the intact ER product release fewer doses at each dosing releasing time point compared to an IR product, which releases the dose at once. Thus, in my opinion, this comparison of ER vs. IR may not be clinically meaningful. We assess abuse potential of ER product by crushing the ER product, and then comparing it with IR product.

2.3.1.2 Study Design

This is a single-dose, randomized, double-blind, triple-dummy, active and placebo controlled crossover, phase I study designed to assess the abuse potential of manipulated IN CEP-33237 in healthy nondependent, recreational opioid users.

The study consisted of 3 phases:
Screening Phase (A): at visit 1, subjects where evaluated to determine if they met criteria for inclusion and exclusion.
Qualification Phase (B): at visit 2, eligible subjects randomized to placebo and hydrocodone powder at dose strength of 45 mg to ensure tolerability and appropriate reporting of positive subjective effects.
Treatment Phase ©: was the randomized, double-blind, triple-dummy, placebo-controlled, 4-period crossover portion of the study. Eligible subjects were randomly assigned to 1 of the following 4 treatment sequences. Each treatment in this phase was separated by a minimum 14-day washout period. The treatment sequences were (ABDC, BCAD, CDBA, and DACB) defined as follows:

- **Treatment A (crushed CEP-33237) -** consisted of 1 intact placebo tablet (matching the 45-mg CEP-33237), 60 mL of a noncarbonated flavored beverage, 1 crushed 45-mg CEP-33237.
- **Treatment B (IR hydrocodone powder) -** consisted of 1 intact placebo tablet (matching the 45-mg CEP-33237), hydrocodone bitartrate powder at a dose strength of 45 mg reconstituted in 60 mL of a noncarbonated flavored beverage, and 1 crushed placebo tablet (matching the 45-mg CEP-33237).
- **Treatment C (intact CEP-33237) -** consisted of 1 intact 45-mg CEP-33237, 60 mL of a noncarbonated flavored beverage, and 1 crushed placebo tablet (matching the 45-mg CEP-33237).
- **Treatment D (placebo)** - consisted of 1 intact placebo tablet (matching the 45-mg CEP-33237), 60 mL of a noncarbonated flavored beverage, and 1 crushed placebo tablet (matching the 45-mg CEP-33237).

2.3.1.3 Abuse Potential Measures

The overall drug liking VAS, the Take Drug Again Assessment (TDAA), and the PVAQ assessment were completed 24 hours after the start of administration of study drug in each period of phase C. The DLEQ was completed at 0.25, 0.75, 1.25, 1.75, 2.5, 4, 6, 7, 8, 9, 10, 12, 24, 36, 48, 60, and 72 hours after the start of administration of study drug in each period of phase C. Pupil diameter measurements were completed prior to study drug administration and at 0.25, 0.75, 1.25, 1.75, 2.5, 4, 6, 7, 8, 9, 10, 12, 24, 36, 48, 60, and 72 hours after the start of administration of study drug in each period of phase C. Questions from the ARCI that comprise the MBG, LSD, and PCAG subscales were completed prior to study drug administration and at 0.25, 0.75, 1.25, 1.75, 2.5, 4, 6, 8, 12, 24, 36, 48, 60, and 72 hours after the start of administration of study drug in each period of phase C.

Primary Endpoints was the Emax of Drug Liking VAS.

The secondary endpoints for assessment of relative abuse potential included

- Drug Liking and Effects Questionnaire (DLEQ)
- Overall Drug Liking Visual Analog Scale (VAS)
- Take Drug Again Assessment (TDAA)
- Addiction Research Center Inventory (ARCI)
- Price Value Assessment Questionnaire (PVAQ)
- Pupil diameter measurement

2.3.1.4 Number of Subject

One hundreds subjects were enrolled. In phase B, data from 97 subjects were analyzed for safety and 92 subjects completed the phase. In the treatment phase, data for 49 subjects were analyzed for safety, data for 45 subjects were analyzed for pharmacodynamics (which received at least 1 dose of study drug) and 35 subjects completed the study (See Figure 19 in Appendix for the detail).

2.3.1.5 Statistical Methodologies used in the Sponsor's Analyses

Continuous and ordinal categorical pharmacodynamic parameters were analyzed using a mixed effects analysis of variance (ANOVA) model that includes study drug, treatment sequence, and period as fixed effects, and subject as a random effect. Comparisons between pairs of treatments were made using the least square means that were estimated from the ANOVA. Pharmacodynamic assessments were also summarized by time point. All analysis will based on the pharmacodynamics analysis set which includes subjects who have adequate pharmacodynamics data from treatment phase to contribute data to at least 1 planned comparison.

The comparison between treatments B and D was assessed first to ensure the validity of the measure. If the treatment difference was significant at an alpha level of 0.05, the comparison between treatments B and C was made. If that treatment difference was also significant at an alpha level of 0.05, the comparison between treatments B and A was made. The following sample SAS code was used to generate the inferential statistics.

```
PROC MIXED DATA=XXXX;
CLASS ARMCD PERIOD TREAT USUBJID;
MODEL Y = ARMCD PERIOD TREAT/S;
RANDOM USUBJID;
LSMEANS TREAT/DIFF=CONTROL('B') PDIFF ALPHA=0.05;
RUN;
```

The sponsor claimed that with 32 evaluable subjects, this study had 90% power to detect a difference of 12 to 20 points on VAS scale (0-100 points) between a pair of treatments using the 2-sided paired t-test with statistical significance of 0.05.

2.3.1.6 Changes in the Conduct of the Study

The Statistical Analysis Plan was finalized on January 10, 2014. The sponsor claimed that the following changes to the analysis sets were made in the SAP prior to database lock and un-blinding. There were no changes to the planned analyses in this study. However, several post hoc analyses were done.

2.3.1.7 Sponsor's Summary and Conclusions

The Sponsor summarized their PD analysis results as follows:

- *Following administration of immediate-release hydrocodone, scores were in the 'liking' range (>50) of the scale between 0.75 and 8.0 hours after administration, after which they returned to just above neutral (50) by 12 hours after administration. Following administration of the extended-release tablet crushed, a slower rise to a lower peak drug liking was observed. Liking scores were generally higher than baseline between 1.25 and 8 hours after administration; however only a small increase in mean scores was observed (approximately 10 points). Drug liking VAS scores following administration of placebo and the extended-release tablet intact had very similar profiles, showing little change across time points and hovering around neutral.*

- *Significantly lower mean Emax of drug liking was observed following administration of the extended-release hydrocodone tablet both intact and crushed, as compared with the immediate-release product, based on Emax, using responses to question 1 of the DLEQ over 72 hours (mean scores of 53.9 and 66.9 vs 85.2; p<0.001, respectively). Mean drug liking VAS Emax for the extended-release tablet intact was not significantly different than placebo (53.9 vs 53.2, p=0.640); however the extended-release tablet crushed was significantly different from placebo (66.9 vs 53.2, p<0.001). Drug liking Emax was also significantly different following administration of the extended-release tablet intact as compared with the extended release tablet crushed (53.9 vs 66.9; p<0.001).*

The sponsor concluded that:

Consistent with the observed differences in pharmacokinetics across treatments, the extended-release hydrocodone product demonstrated significantly lower subjective effects compared to immediate-release hydrocodone, when administered by the oral route as intact or crushed tablets. Substantial reductions were observed with intact extended-release hydrocodone, as it showed a similar profile to placebo. Effects were only slightly less pronounced when the product was crushed. These results demonstrate that this extended-release hydrocodone product may have lower abuse potential compared to non-abuse deterrent opioid products.

2.3.2 Reviewer's Assessment

The sponsor's all analysis were based on the pharmacodynamics analysis set (n=45) which includes subjects who have adequate pharmacodynamics data from phase C to contribute data to at least 1 planned comparison. This review was based on the completer (n=35) in order to be consistent with previous study (C-10032). The results were similar based on these two populations.

2.3.2.1 Primary Endpoint – Drug Liking VAS

Descriptive Statistics show that, following administration of IR hydrocodone, scores were in the 'liking' range (>50) of the scale between 0.75 and 8.0 hours after administration, after which they returned to just above neutral (50) by 12 hours after administration (Figure 7). Following administration of the CEP-33237 tablet crushed, a slower rise to a lower peak drug liking was observed. Liking scores were generally higher than baseline between 1.25 and 8 hours after administration; however only a small increase in mean scores was observed (approximately 10 points). Drug liking VAS scores following administration of placebo and the CEP-33237 tablet intact had very similar profiles, showing little change across time points and hovering around neutral. Figure 8 is the heat map for Drug Liking VAS Emax by treatment for completer population. The density of the color blue indicates the degree of the disliking and the density of the color red indicates the degree of the liking. This figure shows that most of subjects in CEP-33237 doses were neutral and slightly above neutral except there were few subjects liked. The subjects in placebos were neutral and in IN API and IN Zohydro™ doses were highly liked.

Figure 7: The Mean Time Course Profiles on Drug Liking VAS by Treatment (N=35)

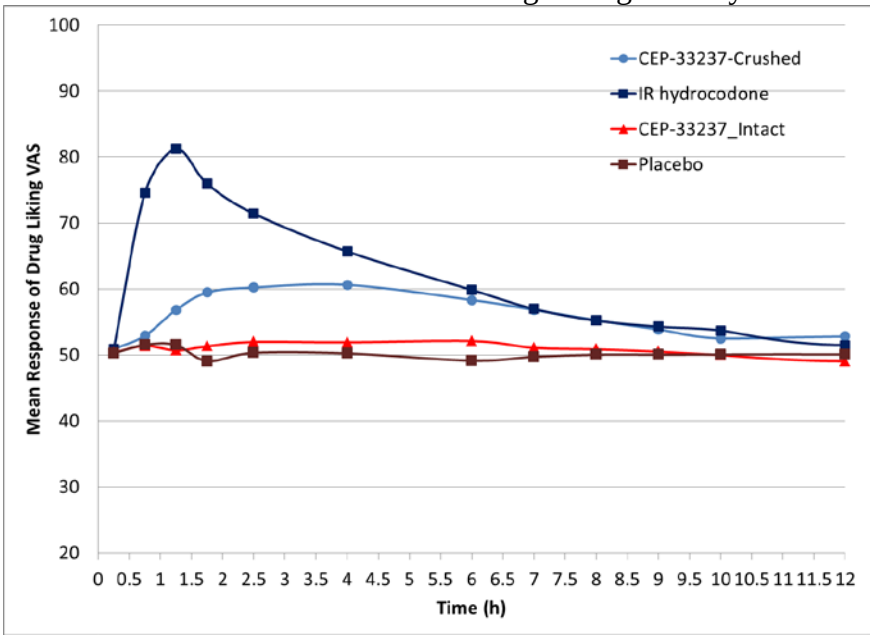


Figure 8: Heat Map for Emax of Drug Liking VAS by Treatment (N=35)

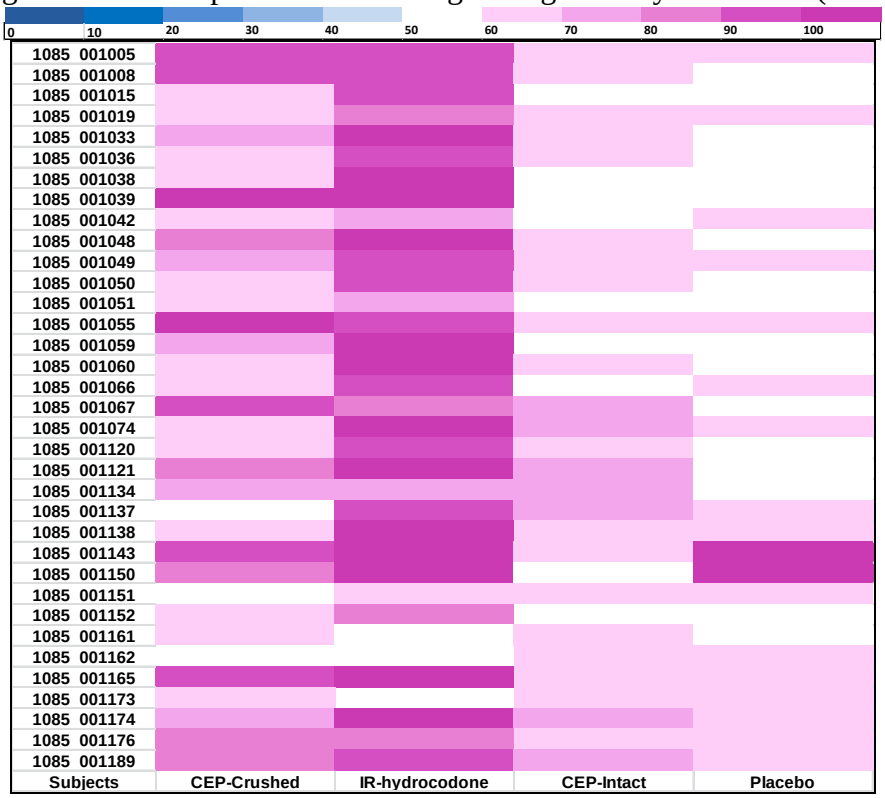


Table 9 shows the first quartiles of the primary endpoint are around 50 for CEP-33237 (crushed), CEP-33237 (intact), and placebo which are much lower than IR hydrocodone (80). Placebo scores remained close to the neutral mark (50), showing very little change across the sampling

period with a mean Emax of 53.1. The mean difference and the median difference between these treatments are about 30 and 40, respectively. As displayed in Table 10, significantly lower mean Emax of drug liking was observed following administration of the CEP-33237 tablet both intact and crushed, as compared with the IR hydrocodone, based on Emax over 72 hours (mean scores of 53.9 and 66.9 vs 85.2; p<0.001, respectively). The comparison with placebo showed that mean drug liking VAS Emax for the CEP-33237 tablet intact was not significantly different than placebo (53.9 vs 53.2, p=0.675); however the CEP-33237 tablet crushed was significantly different from placebo (66.9 vs 53.2, p<0.001). Drug liking Emax was also significantly different following administration of the CEP-33237 tablet intact as compared with the CEP-33237 tablet crushed (53.9 vs 66.9; p<0.001).

Additional analysis of Emax using a nonparametric model (Wilcoxin Signed Rank Test) revealed a similar pattern of results, including statistically significant differences between IR hydrocodone and placebo (median score of 51.0 vs 88.0, p<0.001) and CEP-33237 (intact and crushed) compared with the IR hydrocodone (median scores of 51.0 and 61.5 vs 88.0, p<0.001 for both). In addition, comparisons between placebo and the CEP-33237 showed significantly higher Emax values for the crushed tablet (p<0.001) but not the intact tablet (p=0.086) and CEP-33237 the crushed showed significantly greater scores than the intact (p<0.001). The analysis results for overall drug like are similar to the results of Emax (Table 11).

Table 9: Summary Statistics for Emax of Drug Liking VAS (N=35)

TRTP	N	Mean	StdErr	Min	Q1	Med	Q3	Max
Emax - Drug Liking VAS (DLEQ-Q1)								
ER (crushed)	35	65.60	2.46	50	51	60	79	98
IR hydrocodone	33	84.97	2.31	50	80	88	97	100
ER (intact)	35	54.54	1.02	50	50	51	56	70
Placebo	35	53.37	1.80	50	50	51	51	100
Overall Drug Liking								
ER (crushed)	35	57.57	3.35	3	50	55	74	97
IR hydrocodone	33	74.03	2.97	45	56	77	88	100
ER (intact)	35	51.23	1.21	24	50	50	52	74
Placebo	35	51.11	1.40	35	50	50	50	96

Table 10: The Treatment Comparison in Means of Emax of Drug Liking (VAS) (N=35)

Treatment Comparison	LS Mean Difference	SE	P-Value	Lower of 95% CI	Upper of 95% CI
Emax - Drug Liking VAS (DLEQ-Q1)					
IR – ER (crushed)	19.05	2.53	<.0001	14.02	24.07
IR – ER (intact)	30.11	2.53	<.0001	25.08	35.13
IR - Placebo	31.38	2.53	<.0001	26.36	36.40
ER (crushed) – Placebo	12.34	2.49	<.0001	7.40	17.27
ER (intact) - Placebo	1.27	2.49	0.6095	-3.66	6.21
Overall Drug Liking					
IR – ER (crushed)	16.52	3.30	<.0001	9.98	23.06

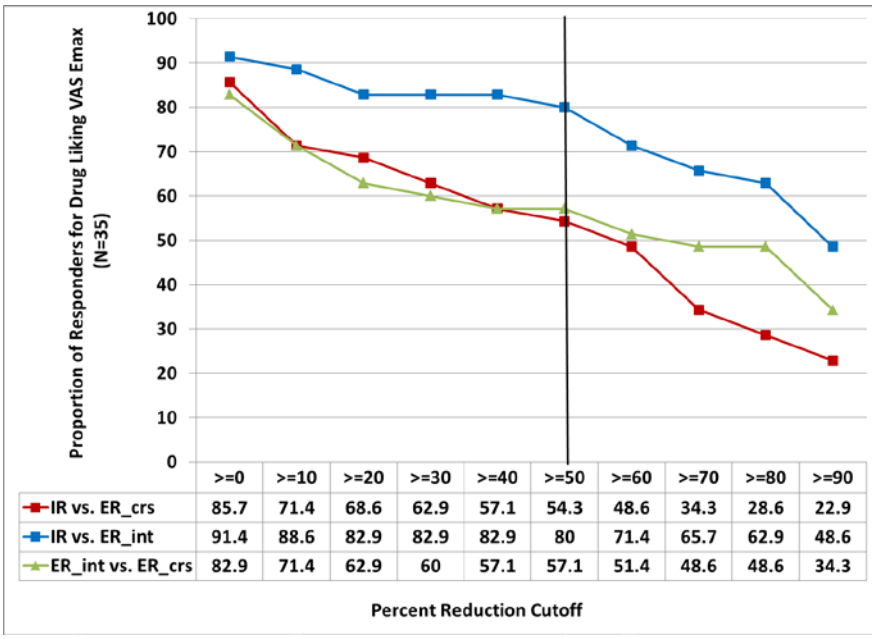
Treatment Comparison	LS Mean Difference	SE	P-Value	Lower of 95% CI	Upper of 95% CI
IR – ER (intact)	22.87	3.30	<.0001	16.33	29.41
IR - Placebo	23.20	3.30	<.0001	16.66	29.74
ER (crushed) – Placebo	6.68	3.30	0.0453	0.14	13.22
ER (intact) - Placebo	0.33	3.30	0.9195	-6.21	6.87

Table 11: The Treatment Comparison in Median of Emax of Drug Liking (VAS) (N=35)

Treatment Comparison	Median Difference	Q1	Q3	P-value
Emax - Drug Liking VAS (DLEQ-Q1)				
IR – ER (crushed)	21	7	30	0.000000
IR – ER (intact)	32	20	41	0.000000
IR - Placebo	33	23	43	0.000000
ER (crushed) – Placebo	8	1	22	0.000002
ER (intact) - Placebo	1	0	6	0.029443
Overall Drug Liking				
IR – ER (crushed)	13	-1	34	0.00034
IR – ER (intact)	24	2	38	0.00000
IR - Placebo	27	6	37	0.00000
ER (crushed) – Placebo	5	0	25	0.02482
ER (intact) - Placebo	0	0	3	0.29746

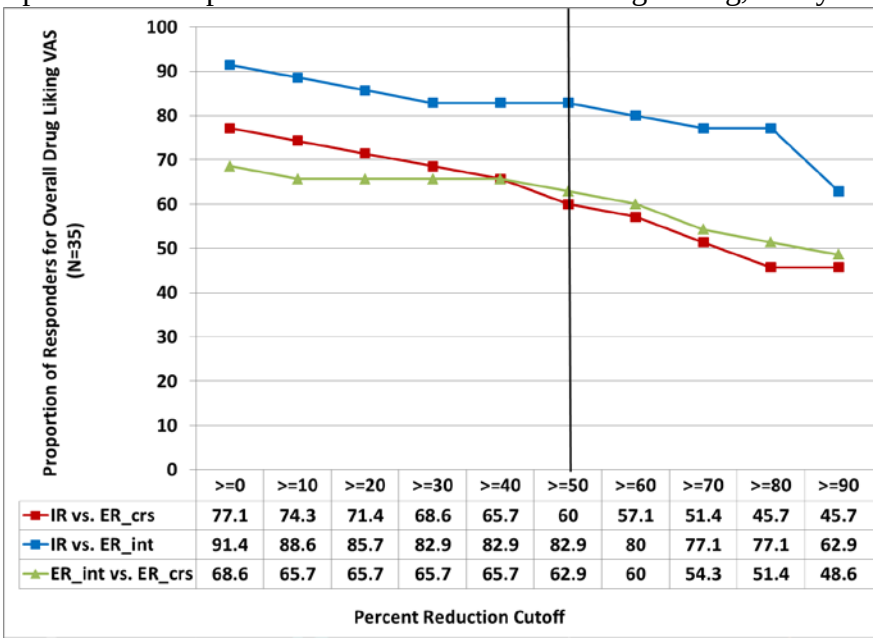
Responder analysis results also support the primary analysis results. Given median of Emax of Drug Liking VAS greater than 88 for IR hydrocodone, approximately 54.3% and 80.0% of subjects had at least 50% reduction for CEP-33237 crushed and intact to IR hydrocodone, respectively (Figure 9). For Overall Drug Liking VAS approximately 60% and 82.9% of subjects had at least 50% reduction for CEP-33237 crushed and intact relative to IR hydrocodone (Figure 10).

Figure 9: Proportion of Responders for Emax of Drug Liking VAS, Study C-1085 (N=35)



	>=0	>=10	>=20	>=30	>=40	>=50	>=60	>=70	>=80	>=90
IR vs. ER_crs	85.7	71.4	68.6	62.9	57.1	54.3	48.6	34.3	28.6	22.9
IR vs. ER_int	91.4	88.6	82.9	82.9	82.9	80	71.4	65.7	62.9	48.6
ER_int vs. ER_crs	82.9	71.4	62.9	60	57.1	57.1	51.4	48.6	48.6	34.3

Figure 10: Proportion of Responders for Emax of Overall Drug Liking, Study C-1085 (N=35)



	>=0	>=10	>=20	>=30	>=40	>=50	>=60	>=70	>=80	>=90
IR vs. ER_crs	77.1	74.3	71.4	68.6	65.7	60	57.1	51.4	45.7	45.7
IR vs. ER_int	91.4	88.6	85.7	82.9	82.9	82.9	80	77.1	77.1	62.9
ER_int vs. ER_crs	68.6	65.7	65.7	65.7	65.7	62.9	60	54.3	51.4	48.6

2.3.2.2 Pupillometry Endpoint

A significant difference in Emin was observed between placebo and IR hydrocodone (5.5 vs 3.2, $p<0.001$). The Emin for pupil diameter measurements was significantly higher (ie, pupils were less constricted) following administration of the CEP-33237 tablet intact or crushed than that following the IR hydrocodone (4.3 vs 3.2 and 4.0 vs 3.2, respectively; $p<0.001$). The Emin following administration of the CEP-33237 tablet intact was also significantly higher than that

for the CEP-33237 tablet crushed (4.3 vs 4.0; p<0.001) (Table 12 and **Error! Reference source not found.**). For all hydrocodone treatments, subjects' pupils were constricted relative to those following placebo administrations. The CEP-33237 tablet intact was associated with a much slower onset compared to the other active treatments and persisted for a longer duration of time. Placebo did not affect pupil diameter (Figure 11).

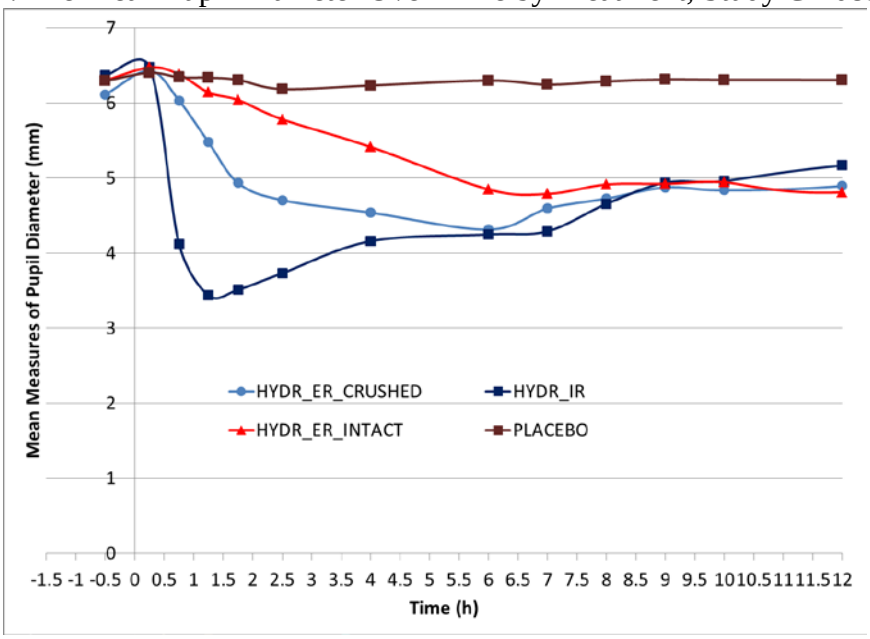
Table 12: Summary Statistics for Emin of Measured Pupillometry (N=35)

Treatment	N	Mean	StdErr	Min	Q1	Med	Q3	Max
Emin, Pupil Diameter								
ER (crushed)	35	3.99	0.13	2.6	3.2	4.1	4.6	5.4
IR	33	3.26	0.10	1.9	2.9	3.2	3.6	4.7
ER (intact)	35	4.33	0.15	2.4	3.7	4.3	5.1	5.9
Placebo	35	5.48	0.17	3.0	4.7	5.6	6.3	7.2

Table 13: Comparison of Minimum Effect as Measured by Pupillometry (N=35)

Treatment Comparison	LS Mean Difference	SE	P-Value	Lower of 95% CI	Upper of 95% CI
Emin, Pupil Diameter					
IR – ER (crushed)	-0.7278	0.1268	<.0001	-0.9795	-0.4762
IR – ER (intact)	-1.0651	0.1268	<.0001	-1.3167	-0.8134
IR - Placebo	-2.2122	0.1266	<.0001	-2.4636	-1.9609
ER (crushed) – Placebo	-1.4844	0.1243	<.0001	-1.7311	-1.2378
ER (intact) - Placebo	-1.1472	0.1243	<.0001	-1.3938	-0.9005

Figure 11: The Mean Pupil Diameter Over Time by Treatment, Study C-1085 (N=35)



2.3.2.3 Conclusion

In this study, a significantly lower mean Emax of drug liking was observed following administration of the CEP 33237 tablet both intact and crushed, as compared with the IR hydrocodone, based on Emax, using responses to Drug Liking VAS over 72 hours (mean scores of 53.9 and 66.9 vs 85.2; $p<0.001$, respectively). The comparison with placebo showed that mean drug liking VAS Emax for the CEP 33237 tablet intact was not significantly different from placebo (53.9 vs 53.2, $p=0.675$); however the CEP 33237 tablet crushed was significantly different from placebo (66.9 vs 53.2, $p<0.001$). Drug liking Emax was also significantly different following administration of the CEP 33237 tablet intact as compared with the CEP 33237 tablet crushed (53.9 vs 66.9; $p<0.001$).

There were the statistical significantly larger pupillary differences between crashed or intact CEP-33237 and placebo in study C-1085.

3 APPENDIX

Table 14: Patients' Demographic and Baseline Characteristics, Study C-10032

Demographic variables	Total (N=45)
Age, years	
n	45
Mean	27.8
SD	8.41
SE	1.25
Median	25.0
Min, max	19.0, 52.0
Sex, n (%)	
Men	33 (73)
Women	12 (27)
Race, n (%)	
White	39 (87)
Black	4 (9)
Asian	1 (2)
Other[a]	1 (2)
Ethnicity, n (%)	
Hispanic or Latino	5 (11)
Non-Hispanic and non-Latino	40 (89)
Weight, kg	
Mean	69.3
SD	11.22
SE	1.67
Median	67.6
Min, max	51.3, 108.0
Height, cm	
Mean	171.1
SD	7.64
SE	1.14
Median	172.0
Min, max	156.0, 193.0
BMI, kg/m^2	
Mean	23.6
SD	3.11
SE	0.46
Median	22.9
Min, max	18.4, 30.4

[Source: Table 7 in C33237-C-10032 study report – report-body.pdf, page 65]

Table 15: Study Procedures and Assessments

Procedures and assessments	Visit 1	Visit 2			Visits 3 through 7					Visit 8
	Phase A	Phase B[a] Drug Discrimination (2 double-blind treatment periods)			Phase C[a] Treatment Phase (5 double-blind, quadruple-dummy treatment periods)					
						48 hour sampling				
	Screening	Check-in (day-1)	Days of study drug administration[b] (day 1 and 3)	Discharge (day 4)	Check-in (day -1)	Days of study drug administration (day 1)	Days 2 and 3	Discharge (day 3, Periods 1-4)	Final assessment (day 3, Period 5) or early withdrawal	Follow-up[c]
C-SSRS	X	X		X	X				X	X
Study restriction compliance review	X	X		X	X			X	X	X
Naloxone challenge		X								
Training for pharmacodynamic assessments		X			X					
Randomization[d]			X			X				
Study drug administration[b]			X			X				
Compliance check			X			X				
DLEQ[f]			X			X	X			
Overall Drug Liking VAS, Take Drug Again VAS and PVAQ[j]			X			X	X			
ARCI (MBG, LSD, and PCAG)[k]			X			X	X			
Ease of Snorting VAS[l]						X				
SRAII[m]						X				
Pupil diameter measurement[n]			X			X	X			
Blood sampling for pharmacokinetics[o]						X	X			
Concomitant medication review[p]		X	X	X	X	X	X	X	X	X
Adverse event inquiry[q]		X	X	X	X	X	X	X	X	X

[Source: Table 1 in report-body.pdf, page 25]

Table 16: Subject Disposition (Randomized Set)

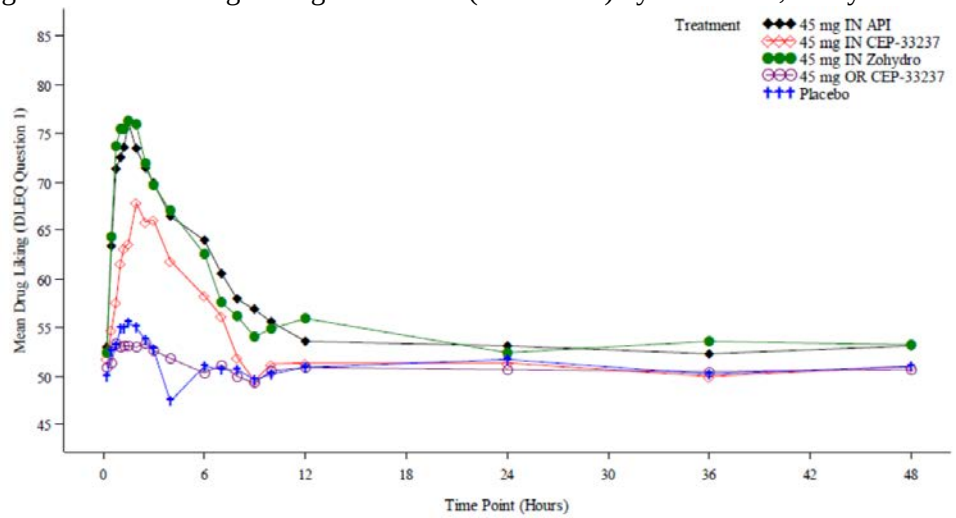
Analysis group, n (%)	Total (N=163)
Screened	163
Screened, not enrolled	90
Adverse event	0
Withdrawal by subject	1
Inclusion criteria not met	50
Exclusion criteria met	36
Non-compliance to study procedures	0
Lost to follow-up	0
Other	3
Enrolled (Randomized in phase B)	73 (100)
Randomized but not treated in phase B	0
Phase B safety analysis set	73 (100)
Completed phase B	71 (97)
Randomized (Randomized in phase C)	45 (62)
Randomized but not treated in phase C	0
Phase C safety analysis set	45 (62)
Pharmacodynamic analysis set	34 (47)
Pharmacokinetic analysis set	42 (58)
Completed study	34 (47)
Discontinued from study	39 (53)
Adverse event	0
Withdrawal by subject	1 (1)
Non-compliance to study medication	1 (1)
Protocol violation	1 (1)

[Source: Summary table 15.1.1 of study report report-body.pdf, page 295]

Table 17: Summary of Analysis Results (p-value) Selected Abuse Potential Measures Study C-10032 (N=34)

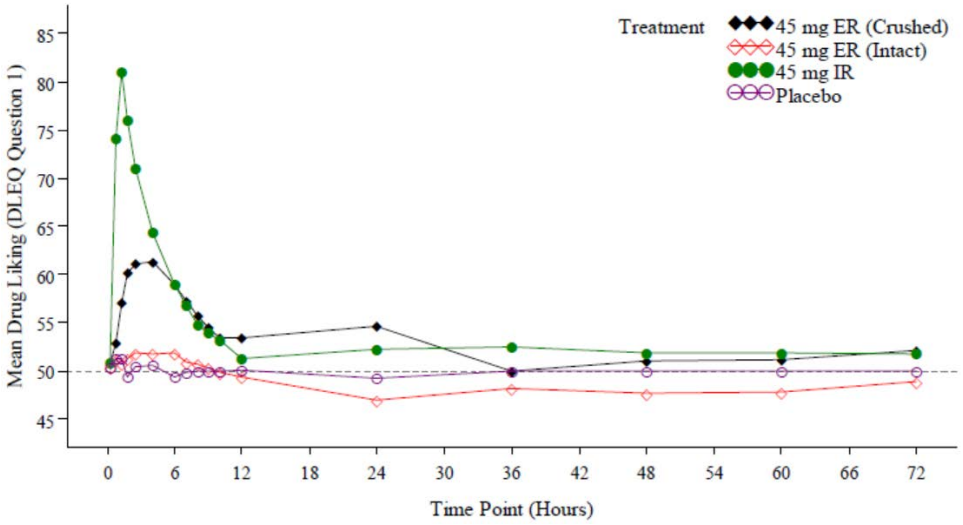
Parameter	IN CEP-33237 vs. IN API	IN CEP 33237 vs. IN Zohydro™	IN Zohydro™ vs. IN API
Primary Endpoints			
Drug Liking VAS E_{max}	**0.004 (↓)**	**<0.001 (↓)**	0.189
Overall Drug Liking VAS E_{max}	**0.004 (↓)**	**<0.001 (↓)**	0.189
Secondary Balance of Effects Endpoints			
Drug Liking VAS E_{min} (Disliking)	**0.0056 (↓)**	**0.0181 (↓)[a]**	0.8388
Overall Drug Liking VAS E_{min} (Disliking)	**0.013 (↓)**	**0.035 (↓)[a]**	0.693
Drug Liking VAS $AUEC_{0-12h}$	**<0.0001 (↓)**	**0.0002 (↓)**	0.9667
Drug Liking VAS $AUEC_{0-48h}$	**<0.0001 (↓)**	**0.0031 (↓)**	0.9002
Take Drug Again VAS E_{max}	**0.005 (↓)**	**<0.001 (↓)**	0.110
PVAQ E_{max}	**0.0342 (↓)**	**0.0002 (↓)**	0.5700
Secondary Positive Effects Endpoints			
Good Effects VAS E_{max}	**<0.0001 (↓)**	**<0.0001 (↓)**	0.0643
Good Effects VAS $AUEC_{0-48h}$	**0.0003 (↓)**	**0.0007 (↓)**	0.9533
ARCI MBG E_{max}	0.1772	0.2998	0.2964
ARCI MBG $AUEC_{0-24h}$	0.4669	0.1613	0.4888
Secondary Negative Effects Endpoints			
Bad Effects VAS E_{max}	**0.0442 (↑)**	0.4187	0.1169
Bad Effects VAS $AUEC_{0-48h}$	0.5783	0.5457	0.8185
Nausea VAS E_{max}	0.8626	0.7457	0.8961
Nausea VAS $AUEC_{0-48h}$	0.4499	0.8489	0.9695
ARCI LSD E_{max}	0.5973	0.1341	0.6323
ARCI LSD $AUEC_{0-24h}$	0.5576	0.9135	0.4348
Secondary Sedative Effects Endpoints			
Alertness/Drowsiness VAS E_{min}	0.073	**0.006 (↓)[a]**	0.337
Alertness/Drowsiness VAS $AUEC_{0-48h}$	0.3162	0.0651	0.4151
ARCI PCAG E_{max}	0.457	0.063	0.259
ARCI PCAG $AUEC_{0-24h}$	0.0651	**0.0408 (↓)**	0.8870
Secondary Other Effects Endpoints			
Any Drug Effects VAS E_{max}	**0.006 (↓)**	**<0.001 (↓)**	0.059
Any Drug Effects VAS $AUEC_{0-48h}$	**0.0043 (↓)**	**0.0024 (↓)**	0.6094
Pupil Diameter E_{min}	0.677	**0.006 (↓)[a]**	**0.019 (↑)[a]**
Pupil Diameter $AUEC_{0-48h}$	0.755	0.487	0.704

Figure 12: Mean Drug Liking Over Time (0-48 hours) by Treatment, Study C-10032



[Source: page 154 in report-body.pdf]

Figure 13: Mean Drug Liking Over Time (0-72 hours) by Treatment, Study C-1085



[Source: page 67 in report-body.pdf]

Figure 14: Time Course Response Profiles for Individual Subjects to the CEP-33237-Intro for Drug Liking VAS (Completers), Study C-10032

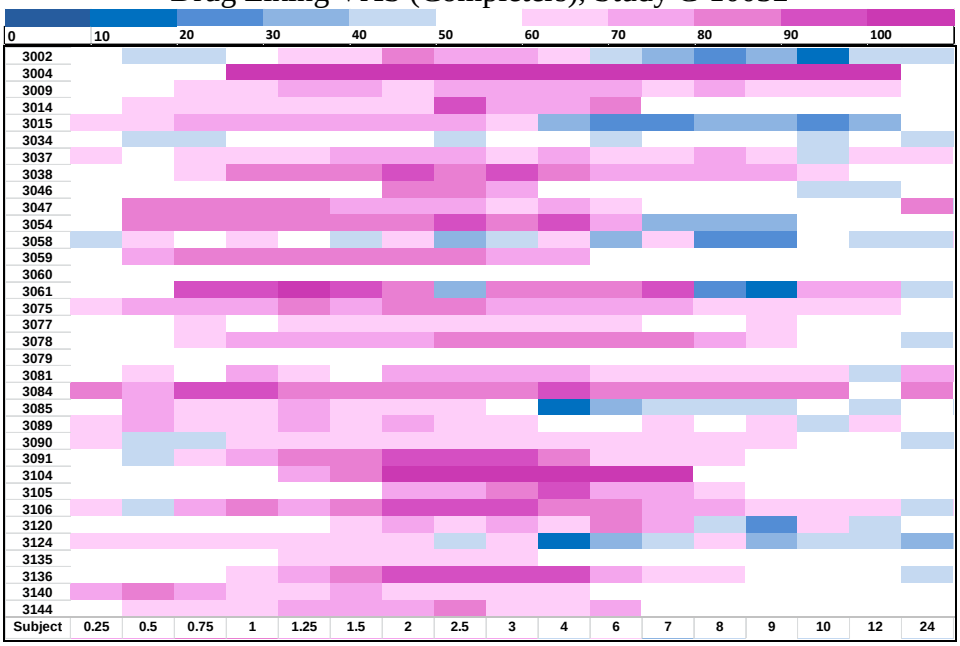


Figure 15: Time Course Response Profiles for Individual Subjects to the Hyd_intro API for Drug Liking VAS (Completers), Study C-10032

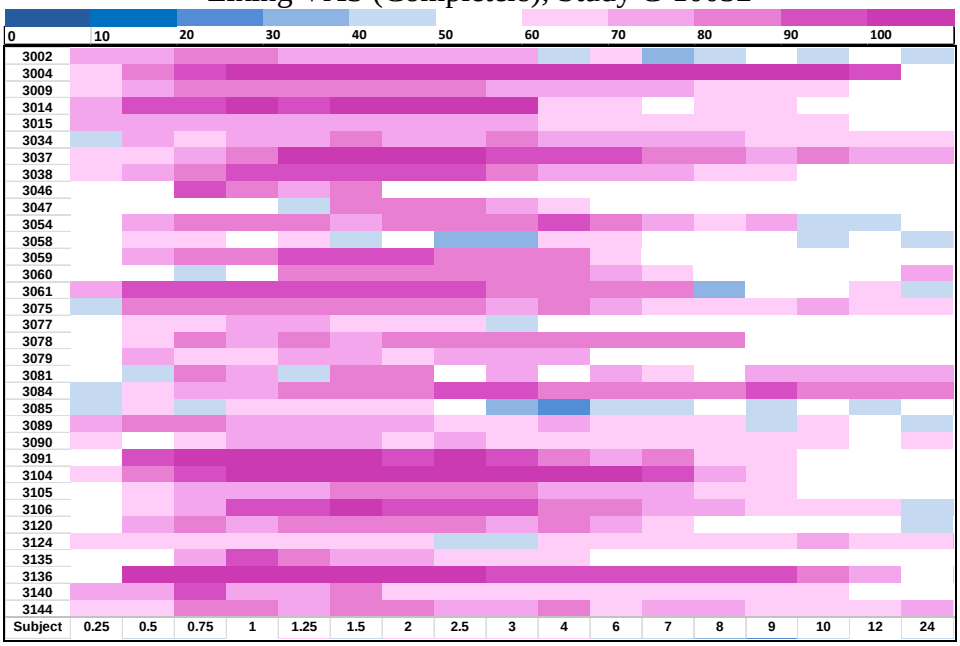


Figure 16: Time Course Response Profiles for Individual Subjects to the CEP-33237_Oral for Drug Liking VAS (Completers), Study C-10032

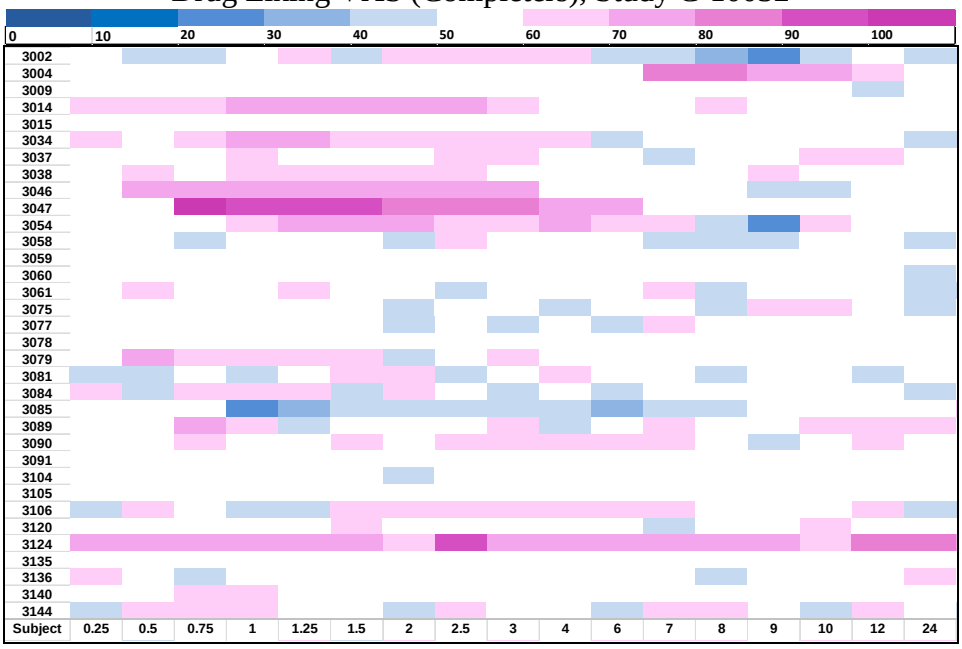


Figure 17: Time Course Response Profiles for Individual Subjects to the Placebo for Drug Liking VAS (Completers), Study C-10032

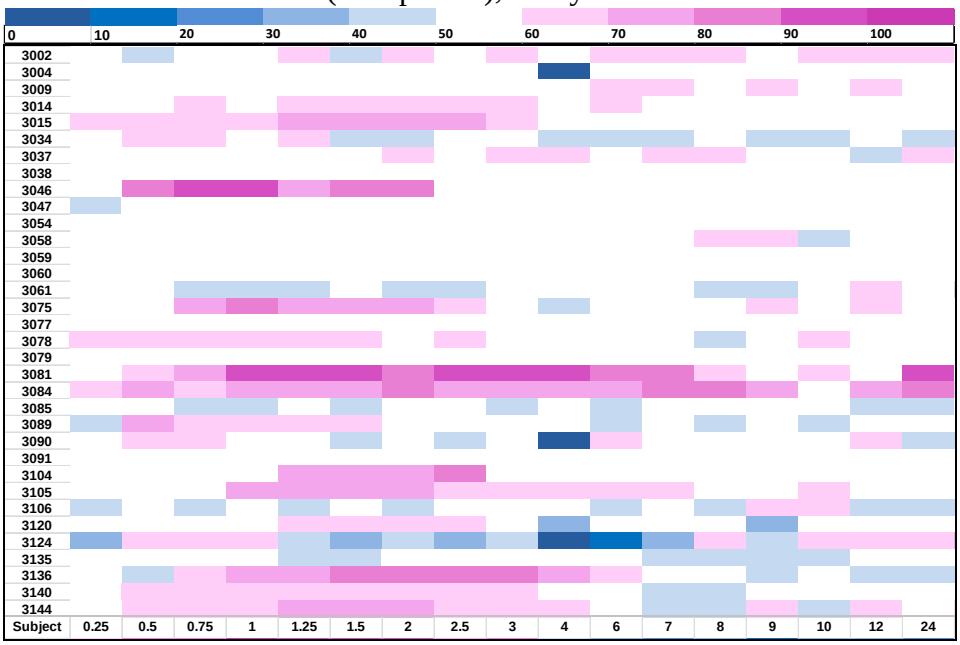


Figure 18: Time Course Response Profiles for Individual Subjects to the Zohydro for Drug Liking VAS (Completers), Study C-10032

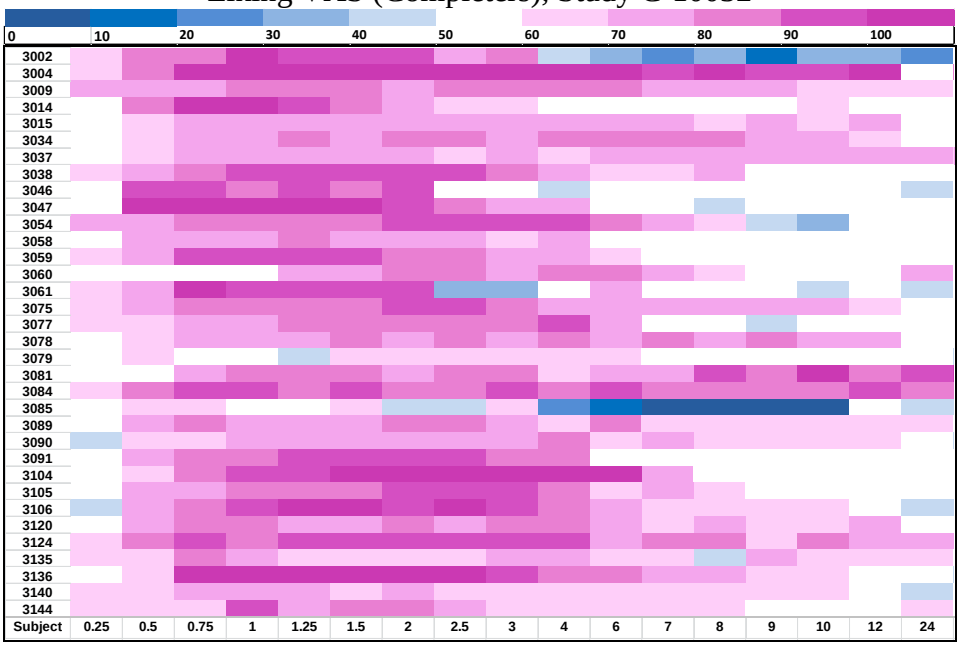
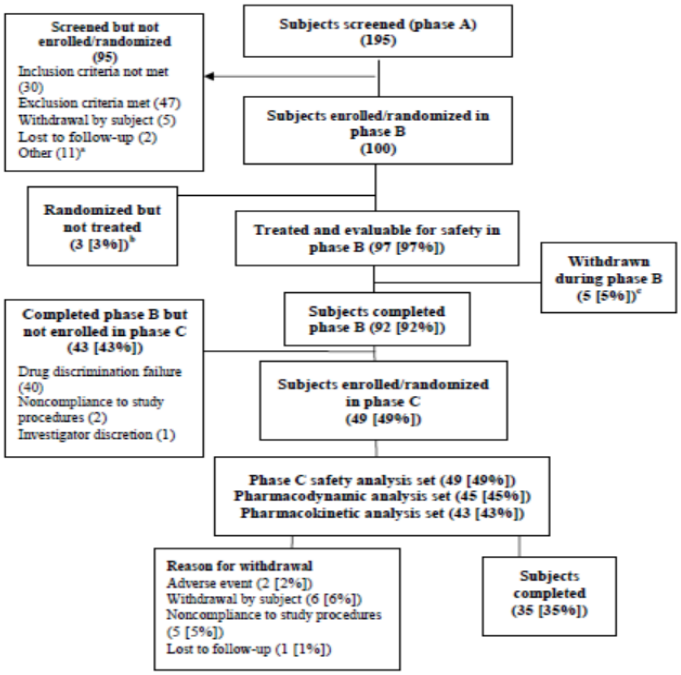


Table 18: Patients' Demographic and Baseline Characteristics, Study C-1085

Demographic variables	Total (N=49)
Age, years	
Mean	24.3
SD	4.84
SE	0.69
Median	23.0
Min, max	18.0, 43.0
Sex, n (%)	
Men	39 (80)
Women	10 (20)
Race, n (%)	
White	46 (94)
Black	1 (2)
Native Hawaiian or other Pacific Islander	1 (2)
Other[a]	1 (2)
Ethnicity, n (%)	
Hispanic or Latino	3 (6)
Non-Hispanic and non-Latino	46 (94)
Weight, kg	
Mean	76.3
SD	12.58
SE	1.80
Median	73.2
Min, max	56.2, 112.0
Height, cm	
Mean	176.0
SD	8.49
SE	1.21
Median	177.8
Min, max	157.5, 190.5
BMI, kg/m^2	
Mean	24.5
SD	3.12
SE	0.45
Median	23.8
Min, max	19.5, 32.6

[Source: Table 7 in C33237/1085 study report – report-body.pdf, page 58]

Figure 19: Subject Disposition



[Source: page 56 in report-body.pdf]

Table 19: The Sponsor's Analysis Results for Drug Liking VAS Emax, Emin, and TA-AUE for Study C-1085 (Completers)

	E_{max}			E_{min}			TA_AUE		
	Median Difference	IQR	P Value	LS Mean Difference (SE)	95% CI	P Value	Median Difference	IQR	P Value
Overall Treatment Effect	---	---	**<0.0001**	---	---	**<0.0001**	---	---	**<0.0001**
Pairwise Comparisons									
Oxycodone HCl IR 15 mg – Placebo lactose	36.0	16.5, 49.0	**<0.0001**	2.7 (4.18)	-5.5, 11.0	0.5186	3.68	0.80, 12.4	**<0.0001**
Oxycodone HCl IR 30 mg – Placebo lactose	47.0	28.0, 49.0	**<0.0001**	0.1 (4.18)	-8.2, 8.3	0.9881	4.85	2.05, 21.5	**<0.0001**
Eluxadoline 100 mg – Oxycodone HCl IR 15 mg	-28.0	-48.0, -12.0	**<0.0001**	-31.1 (4.18)	-39.4, -22.9	**<0.0001**	-8.02	-26.8, -1.65	**<0.0001**
Eluxadoline 200 mg – Oxycodone HCl IR 15 mg	-26.0	-48.0, -9.0	**<0.0001**	-35.7 (4.20)	-44.0, -27.4	**<0.0001**	-9.09	-29.9, -4.18	**<0.0001**
Eluxadoline 100 mg – Oxycodone HCl IR 30 mg	-45.0	-49.0, -26.0	**<0.0001**	-28.5 (4.19)	-36.8, -20.2	**<0.0001**	-10.3	-36.2, -3.41	**<0.0001**
Eluxadoline 200 mg – Oxycodone HCl IR 30 mg	-42.0	-49.0, -11.0	**<0.0001**	-33.1 (4.18)	-41.4, -24.8	**<0.0001**	-8.83	-56.4, -5.08	**<0.0001**
Eluxadoline 100 mg – Placebo eluxadoline	0.0	-1.0, 1.0	0.7352	-21.7 (4.12)	-29.9, -13.6	**<0.0001**	-2.53	-13.3, 0.15	**0.0010**
Eluxadoline 200 mg – Placebo eluxadoline	0.0	0.0, 12.0	0.2500	-26.3 (4.14)	-34.5, -18.1	**<0.0001**	-3.71	-19.0, -0.58	**0.0002**
Eluxadoline 100 mg – Placebo lactose	0.0	-1.0, 2.0	0.4102	-28.4 (4.18)	-36.7, -20.2	**<0.0001**	-1.20	-13.9, 0.04	**0.0007**
Eluxadoline 200 mg – Placebo lactose	0.0	-1.0, 20.0	0.2174	-33.0 (4.20)	-41.3, -24.7	**<0.0001**	-4.37	-19.0, -0.35	**0.0002**
Placebo eluxadoline – Placebo lactose	0.0	0.0, 0.0	0.7698	-6.7 (4.13)	-14.9, 1.4	0.1058	0.00	-0.25, 0.22	0.5740

[Source: Table 11 in cps-1010-study-report.pdf; page 68/3893]

Table 20: Patients' Demographic and Baseline Characteristics, Study C-1085

Demographic variables	Total (N=49)
Age, years	
Mean	24.3
SD	4.84
SE	0.69
Median	23.0
Min, max	18.0, 43.0
Sex, n (%)	
Men	39 (80)
Women	10 (20)
Race, n (%)	
White	46 (94)
Black	1 (2)
Native Hawaiian or other Pacific Islander	1 (2)
Other[a]	1 (2)
Ethnicity, n (%)	
Hispanic or Latino	3 (6)
Non-Hispanic and non-Latino	46 (94)
Weight, kg	
Mean	76.3
SD	12.58
SE	1.80
Median	73.2
Min, max	56.2, 112.0
Height, cm	
Mean	176.0
SD	8.49
SE	1.21
Median	177.8
Min, max	157.5, 190.5
BMI, kg/m^2	
Mean	24.5
SD	3.12
SE	0.45
Median	23.8
Min, max	19.5, 32.6

[Source: Table 5 in report-body.pdf, page 58]

Table 21: Summary Statistics for Other Abuse Potential Measure, Study C-1085

Variable Statistic	Placebo (N=42)	45-mg IR (N=39)	45-mg ER crushed (N=42)	45-mg ER intact (N=41)
DLEQ Q1 E_{min}				
Median	49.5	49.0	49.0	50.0
Mean (SD)	45.3 (12.95)	46.8 (9.77)	46.6 (9.70)	46.2 (10.64)
Difference from IR	−1.50	NC	−0.12	−0.66
95% confidence interval	−6.31, 3.30	–	−4.91, 4.68	−5.50, 4.18
p-value	0.537	–	0.962	0.786
Difference from ER crushed	NC	NC	NC	−0.55
95% confidence interval	–	–	–	−5.29, 4.19
p-value	–	–	–	0.819
DLEQ Q1 AUEC				
Median	3600	3722	3636	3602
Mean (SD)	3595 (110)	3860 (585)	3803 (592)	3544 (308)
Difference from IR	−285	NA	−70	−324
95% confidence interval	−464, −105	–	−249, 109	−505, −144
p-value	0.002[a]	–	0.442	<0.001[a]
Difference from ER crushed	NC	NC	NA	−254
95% confidence interval	–	–	–	−431, −78
p-value	–	–	–	0.005[a]
TDAA				
Median	50.0	74.0	56.0	50.0
Mean (SD)	47.2 (15.51)	75.2 (17.27)	58.7 (21.47)	46.4 (18.27)
Difference from IR	−28.18	NA	−16.92	−28.34
95% confidence interval	−35.96, −20.40	–	−24.69, −9.14	−36.18, −20.51
p-value	<0.001[a]	–	<0.001[a]	<0.001[a]
Difference from ER crushed	NC	NC	NA	−11.43
95% confidence interval	–	–	–	−19.17, −3.69
p-value	–	–	–	0.004[a]
PVAQ				
Median	0.0	10.0	5.0	0.0
Mean (SD)	0.7 (3.38)	12.1 (7.98)	7.3 (11.25)	2.3 (4.84)
Difference from IR	−11.34	NA	−4.78	−9.68
95% confidence interval	−14.21, −8.48	–	−7.64, −1.92	−12.57, −6.79
p-value	<0.001[a]	–	0.001[a]	<0.001[a]
Difference from ER crushed	NC	NC	NA	−4.91
95% confidence interval	–	–	–	−7.76, −2.05
p-value	–	–	–	<0.001[a]

[Source: Table 10 in report-body.pdf, page 73]

Figure 20: Time Course Response Profiles for Individual Subjects to the Hydr ER (crushed) for Drug Liking VAS (Completers), Study C-1085

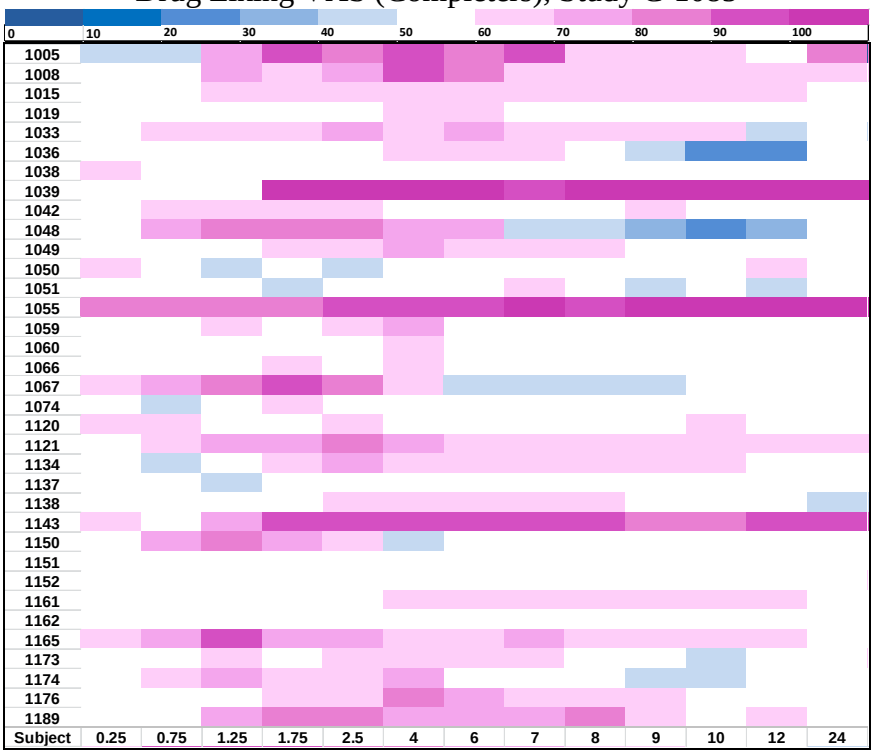


Figure 21: Time Course Response Profiles for Individual Subjects to the Hydr IR for Drug Liking VAS (Completers), Study C-1085

Figure 22: Time Course Response Profiles for Individual Subjects to the Hydr ER (intact) for Drug Liking VAS (Completers), Study C-1085

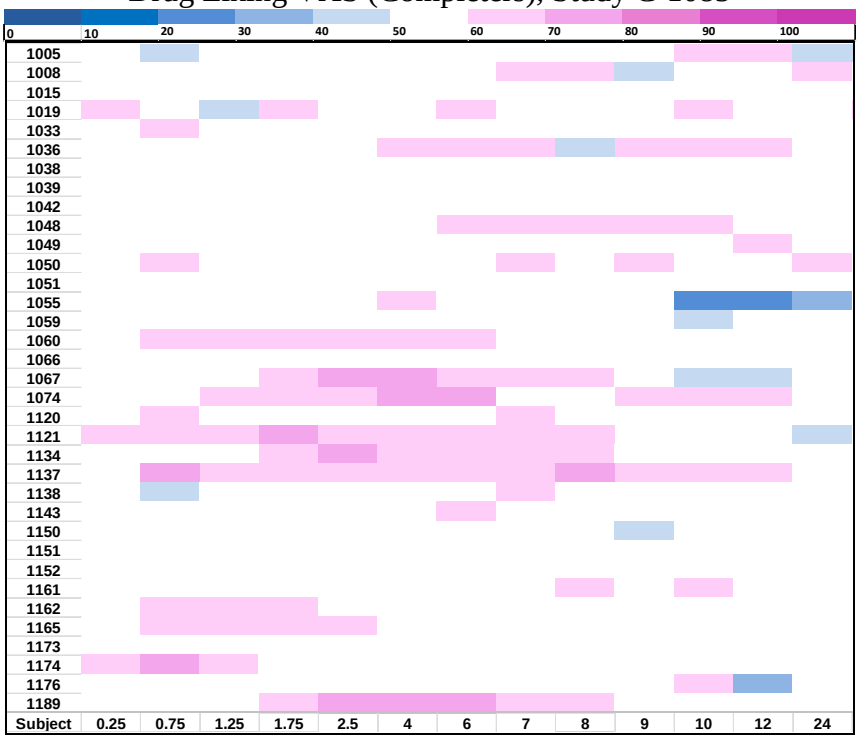
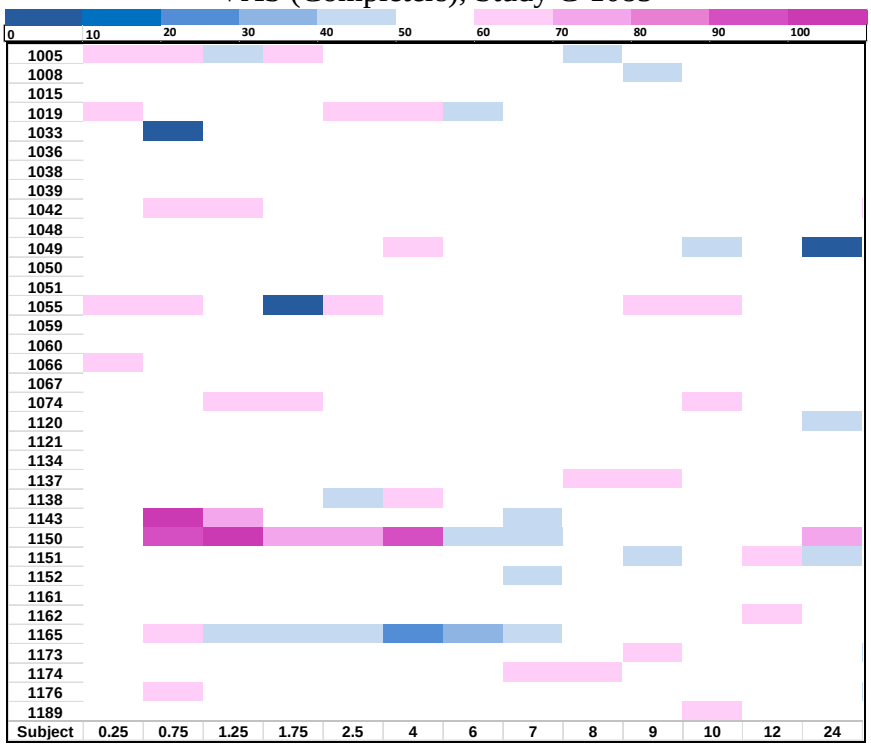


Figure 23: Time Course Response Profiles for Individual Subjects to Placebo for Drug Liking VAS (Completers), Study C-1085



--EOF--

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

FENG ZHOU
09/09/2015

QIANYU DANG
09/09/2015

YI TSONG
09/09/2015

STATISTICS FILING CHECKLIST

NDA Number: 207-975	**Applicant: Teva Branded Pharm. Products R&D, Inc.**	**Stamp Date: Dec 23, 2014**
Drug Name: Hydrocodone bitartrate extended-release, abuse-deterrent tablet	**NDA Type: Standard**	**Indication: Management of pain severe enough to require daily, around-the-clock, long-term opioid treatment and for which alternative treatment options are inadequate**

On **initial** overview of the Supplemental NDA application for RTF:

	Content Parameter for RTF	Yes	No	NA	Comments
1	Electronic Submission: Indexing and reference links within the electronic submission are sufficient to permit navigation through the submission, including access to reports, tables, data, etc.	X			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	X			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated.	X			
4	Data sets in EDR are accessible and conform to applicable guidances (e.g., existence of define.pdf file for data sets).	X			

THE STATISTICAL SECTION OF THE APPLICATION IS FILEABLE

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	X			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	X			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.			X	Not present
Appropriate references for novel statistical methodology (if present) are included.			X	Not present
Safety data organized to permit analyses across clinical trials in the NDA.	X			
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.	X			

STATISTICS FILING CHECKLIST

BRIEF SUMMARY OF CONTROLLED PHASE 3 CLINICAL TRIALS

Study Number (Dates Conducted)	Number of Sites	Sample Size	Type of Control	Design	Duration of Treatment
C33237/3103 (03/2013 – 02/2014)	78 sites in US	Enrolled/screened: 625 Randomization: Hydrocodone: 191 Placebo: 180	Placebo	Multicenter, randomized, double-blind, placebo-controlled, randomized-withdrawal study in patients with moderate to severe chronic low back pain	Open-label titration period: up to 6 weeks Randomization period: 12 weeks
C33237/3079 (11/2010 – 08/2011)	71 sites in US	Enrolled/screened: 391 Randomization: Hydrocodone: 146 Placebo: 148	Placebo	Multicenter, randomized, double-blind, placebo-controlled, randomized-withdrawal study in patients with moderate to severe pain associated with osteoarthritis or low back pain	Open-label titration period: up to 6 weeks Randomization period: 12 weeks

---

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

FREDA COONER
02/19/2015

THOMAS J PERMUTT
02/19/2015
I concur with the conclusion that the submission is filable from the point of view of statistics.