CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

22580Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Pharmacoepidemiology and Statistical Science
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	NDA 22-580/SN000
Drug Name:	Qnexa (phentermine and topiramate)
Indication(s):	Treatment of obesity in patients with and without weight-related comorbidities
Applicant:	Vivus, Inc.
Date(s):	Goal date: 3/24/2012
Review Priority:	Standard
Biometrics Division:	Division VII
Statistical Reviewer:	Ben Neustifter, Ph.D.
Concurring Reviewers:	Mat Soukup, Ph.D., Team Lead, DB VII
	Aloka Chakravarty, Division Director, DB VII
Medical Division:	Division of Metabolism and Endocrinology
Clinical Team:	Eric Colman, M.D., Clinical Team Leader, DMEP
	Mary Roberts, M.D., Clinical Reviewer, DMEP
Project Manager:	Pooja Dharia, Regulatory Project Manager, DMEP

Keywords: Cardiovascular, MACE, Cox regression, weight, weight loss, obesity

Table of Contents

LIST OF TABLES **3**

LIST OF FIGURES **4**

1. EXECUTIVE SUMMARY **5**

1.1 CONCLUSIONS AND RECOMMENDATIONS 5
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES 5
1.3 STATISTICAL ISSUES AND FINDINGS 6

2. INTRODUCTION **6**

2.1 OVERVIEW 6
2.2 DATA SOURCES 7

3. STATISTICAL EVALUATION **7**

3.1 EVALUATION OF EFFICACY 7
3.2 EVALUATION OF SAFETY 7
3.2.1 Analysis of MACEs *7*
3.2.2. Summaries of Heart Rate, Systolic Blood Pressure, Diastolic Blood Pressure, and Rate Pressure Product *10*

5. SUMMARY AND CONCLUSIONS **12**

5.1 CONCLUSIONS AND RECOMMENDATIONS 12

APPENDIX I **13**

APPENDIX II **16**

SIGNATURES/DISTRIBUTION LIST **20**

LIST OF TABLES

Table 1: MACE Hazard Ratios vs. Placebo, Stratified by Study....8
Table 2: MACE Hazard Ratios vs. Placebo, Stratified by Study, OB-305 Excluded....8
Table 3: MACE Hazard Ratios vs. Placebo, No Stratification....9
Table 4: MACE Hazard Ratios vs. Placebo, No Stratification, OB-305 Data Excluded....9
Table 5: Sponsor-provided summary of cardiovascular health markers for one-year cohort....13
Table 6: Sponsor-provided summary of cardiovascular health markers for two-year cohort....14
Table 7: Sponsor-provided summary of Rate Pressure Product for one-year cohort....15

LIST OF FIGURES

Figure 1: Hazard Ratios and Confidence Intervals for MACEs by Study 10
Figure 2: Mean change in Heart Rate for OB-302 and OB-303 Completers 16
Figure 3: Mean change in Systolic Blood Pressure for OB-302 and OB-303 Completers 17
Figure 4: Mean change in Diastolic Blood Pressure for OB-302 and OB-303 Completers.....17
Figure 5: Mean change in Rate Pressure Product for OB-302 and OB-303 Completers 18
Figure 6: Mean change in Heart Rate for OB-303/305 Completers 18
Figure 7: Mean change in Systolic Blood Pressure for OB-303/305 Completers.....19
Figure 8: Mean change in Diastolic Blood Pressure for OB-303/305 Completers 19
Figure 9: Mean change in Rate Pressure Product for OB-303/305 Completers 20

1. EXECUTIVE SUMMARY

Qnexa (herein referred to as PHEN/TPM) is a fixed-dose combination of immediate-release phentermine (PHEN) and modified-release topiramate (TPM), submitted by the sponsor, Vivus Inc., indicated for weight management in obese (BMI $\geq$ 30 kg/m^2) and overweight (BMI 27-29.9 kg/m^2) individuals when accompanied by at least one weight-related comorbidity. PHEN/TPM was studied and has been investigated in Phase 2 and 3 studies using three dosage strengths: low (3.75mg PHEN/23 mg TPM), mid (7.5mg PHEN/46mg TPM), and high (15mg PHEN/92mg TPM).

The Division of Metabolism and Endocrinology submitted a consultation request to Office of Biostatistics, Division VII for a quantitative analysis of cardiovascular-related safety issues. This includes inferential analyses of Major Adverse Cardiac Events (MACEs), and descriptive analyses of heart rate, systolic and diastolic blood pressure, and rate pressure product (RPP).

1.1 Conclusions and Recommendations

Overall, an analysis of the data from studies OB-202, OB-230, OB-302, OB-303, and OB-305 showed no notable differences in the hazard ratio of MACEs between any dose of PHEN/TPM and placebo, for several different methods of stratification and treatment regiments.

Descriptive analyses, including graphical representations, of systolic and diastolic blood pressure, and RPP did not, for the most part, appear to show any significant increases from baseline to the end of study. The mean change at Weeks 56 and 108 in PHEN/TPM-exposed subjects for these three cardiovascular health markers tended to show a slight decrease (possibly indicative of improved cardiovascular health) that was similar to or stronger than that observed in the placebo arm.

For heart rate, however, a significant difference in mean change from baseline after one year was found in the high-dose arm above the placebo. A similar increase from baseline was shown for the two-year data. While increases in heart rate at one year were also shown for the low-dose and the mid-dose, these increases were not significant, and may not indicate a safety risk. It is worth noting, however, that the low and mid-dose had smaller sample sizes (234 and 488, respectively) than the high-dose (1553), and that the lack of significance of these results may be related to the reduced sample sizes.

1.2 Brief Overview of Clinical Studies

OB-202 was a placebo-controlled Phase 2 study of efficacy and safety consisting of 28 weeks of high-dose PHEN/TPM treatment. DM-230 was a 28-week extension of all sites and eligible subjects that wished to continue from trial OB-202; subjects received high-dose PHEN/TPM, unless tolerability caused a reduction during OB-202 or DM-230, in which case they were assigned a mid-dose. These Phase 2 studies are included for the calculation of MACE risk, due to the rarity of MACEs and the subsequent lack of power in a MACE analysis; although the enrollment numbers for OB-202/DM-230 are smaller than those of the Phase 3 studies, the

addition of a year's worth of data to the calculation of MACE risk provides important exposure information to the calculation.

OB-302 and OB-303 were placebo-controlled studies of 1-year exposure to PHEN/TPM. OB-302 enrolled obese adults (BMI $\geq$ 35 kg/m^2) with limited weight-related co-morbidities, and OB-303 enrolled overweight and obese adults (BMI $\geq$ 30 kg/m^2 and $\leq$ 45 30 kg/m^2) with weight-related comorbidities, including diabetes. OB-305 was a 1-year extension study of eligible patients from *applicant-selected* sites from Study OB-303. The selection of sites for extension to study OB-305 was not random; the sponsor has stated that the selection was based on the number of eligible subjects and that it was made while study OB-303 was ongoing. Of the 2,487 subjects randomized in OB-303, 866 (34.8%) were eligible to enroll in the OB-305 extension, and 677 (27.2% of those randomized in OB-303, 78.2% of those eligible for OB-305) enrolled. Since site selection was non-random and treatments were not re-randomized for OB-305, results that utilize this extension's data should be interpreted with caution.

1.3 Statistical Issues and Findings

Cox regression analysis under a variety of conditions did not show a significant increase in MACE risk for any dose of PHEN/TPM (Tables 1 through 4, Figure 1).

As stated in the conclusions, heart rate showed a significant difference in change from baseline after one year in the high-dose arm (mean increase of 1.6 bpm, 95% CI from 1.09 to 2.11) above the placebo (mean change of 0.0 bpm, 95% CI from -0.51 to 0.51), with a similar increase from baseline at two years (mean increase of 1.7, 95% CI from 0.37 to 3.03). Increases in heart rate at one year were also shown for the low-dose (1.3bpm mean, 95% CI from -0.03 to 2.63) and the mid-dose (0.6bpm mean increase, 95% CI from -0.31 to 1.51), but these increases were not significant, and may not indicate a safety risk; however, these comparisons may be underpowered due to small sample size.

Furthermore, descriptive and graphical representations of data from OB-302 and OB-303/305 suggest that the favorable decrease in systolic blood pressure, diastolic blood pressure, and RPP at one year may be moderated during the second year of treatment, as the mean improvement in these values was reduced somewhat during the second year of PHEN/TPM treatment. The clinical significance of these results is a matter for clinical consideration.

2. INTRODUCTION

2.1 Overview

Phentermine (PHEN) was approved in 1959 for the treatment of obesity, and has been indicated for short-term use only since 1973. Topiramate (TPM) was approved for the treatment of seizures in 1996 and for the prevention of migraine headache in 2004. Both are currently licensed for use in the U.S.

On December 28, 2009, the applicant, VIVUS, Inc. submitted a New Drug Application (NDA) seeking approval of PHEN/TPM for the treatment of obesity.

On October 28, 2010, the Agency provided a Complete Response (CR) to the sponsor's application for PHEN/TPM. The two reasons cited for the CR were: animal data and preliminary data from the North American Antiepileptic Drug Pregnancy Registry that suggest TPM poses teratogenic risk to women of child-bearing potential; and cardiovascular safety concerns based on a larger percentage of subjects in Phase 3 trials developing increases in heart rate in PHEN/TPM compared with placebo.

The sponsor resubmitted the NDA on October 17, 2011; their resubmission included data and analyses of Phase 2 and Phase 3 studies of PHEN/TPM, as well as a one-year extension of a Phase 3 study in sponsor selected sites. It should be noted that the one-year extension study is the only new data submitted from that provided in the original submission. As the Division of Biometrics 7 was not consulted to review the cardiovascular safety data in the original submission this review incorporates all trial information in the assessment of cardiovascular safety.

2.2 Data Sources

The sponsor reports and data sets used in the evaluation of this NDA from trials OB-202, DM-230, OB-302, OB-303, and OB-305 are located at the eCTD location: \\cdsesub1\EVSPROD\NDA022580\022580.enx.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

This review does not include any efficacy evaluation.

3.2 Evaluation of Safety

3.2.1 Analysis of MACEs

Major Adverse Cardiac Events (MACEs) were of particular safety concern for PHEN/TPM. MACEs, by their strictest definition, are made up of the following events: cardiovascular-related death, non-fatal myocardial infarction, and non-fatal stroke. The sponsor submitted cardiovascular event reports, including MACE analyses; these were replicated and expanded upon for the safety analysis provided here.

The primary MACE analysis utilized data from OB-202/DM-230, OB-302, OB-303, and OB-305 for a time-to-event analysis in the intent-to-treat (ITT) populations. The ITT populations consisted of all subjects that took at least one dose of their randomized treatment. Due to the above-stated concerns regarding the selection and randomization loss for the OB-305 population, this analysis was repeated without the data from OB-305; OB-303/305 subjects had their data censored at the end of OB-303 for this analysis.

Analysis was performed using a Cox proportional hazard model under a variety of specifications; these analyses were not pre-specified, and so several different models were implemented to check for the robustness of results. The primary model was stratified by study, treating OB-202/DM-230 as a single study (since all eligible subjects were allowed to participate in DM-230), but treating OB-303 subjects as coming from a separate study as OB-303/305 subjects due to the randomization concerns. Therefore, the primary model had four stratification levels: OB-202/DM-230, OB-302, OB-303, and OB-303/305. Furthermore, since sample sizes in some arms were low and the hazard ratios for MACEs by dose were not statistically significantly different, the analysis was repeated with a pooled treatment arm that did not separate by dose. Table 1 provides the Hazard Ratios (HRs) and associated 95% confidence intervals for MACEs in the treatment arms versus placebo.

Table 1: MACE Hazard Ratios vs. Placebo, Stratified by Study

	Placebo	**PHEN/TPM Low (3.75mg/23mg)**	**PHEN/TPM Mid (7.5mg/46mg)**	**PHEN/TPM High (15mg/92mg)**
N	1611	240	498	1607
MACE (%)	4 (0.25%)	1 (0.42%)	2 (0.40%)	4 (0.25%)
HR (95% CI)	-	*	0.99 (0.18, 5.53)	0.90 (0.22, 3.60)
	Pooled PHEN/TPM: HR=1.09 and 95% CI = (0.32, 3.73)			

* As low-dose PHEN/TPM was only used in OB-302, HR estimates for this dose when stratified by study are overly large and difficult to interpret, so are not included.

Since there were randomization concerns with the OB-305 extension to study OB-303, this analysis was repeated excluding the OB-305 data. For this analysis, then, there were only three stratification levels for study: OB-202/DM-230, OB-302, and OB-303. Table 2 provides the HRs and confidence intervals for MACEs the stratified analysis without OB-305 data.

Table 2: MACE Hazard Ratios vs. Placebo, Stratified by Study, OB-305 Excluded

	Placebo	**PHEN/TPM Low (3.75mg/23mg)**	**PHEN/TPM Mid (7.5mg/46mg)**	**PHEN/TPM High (15mg/92mg)**
N	1611	240	498	1607
MACE (%)	3 (0.18%)	1 (0.42%)	1 (0.20%)	3 (0.19%)
HR (95% CI)	-	*	0.70 (0.07, 6.91)	0.92 (0.19, 4.54)
	Pooled PHEN/TPM: HR=1.07 and 95% CI = (0.25, 4.48)			

* As low-dose PHEN/TPM was only used in OB-302, HR estimates for this dose when stratified by study are overly large and difficult to interpret, so are not included.

To assess for robustness of results, secondary models that did not stratify by study were also used. Table 3 provides the HRs and confidence intervals for MACEs for an unstratified analysis that includes data from OB-305.

Table 3: MACE Hazard Ratios vs. Placebo, No Stratification

	Placebo	**PHEN/TPM Low (3.75mg/23mg)**	**PHEN/TPM Mid (7.5mg/46mg)**	**PHEN/TPM High (15mg/92mg)**
N	1611	240	498	1607
MACE (%)	4 (0.25%)	1 (0.42%)	2 (0.40%)	4 (0.25%)
HR (95% CI)	-	2.87 (0.30, 27.60)	1.04 (0.19, 5.73)	0.87 (0.22, 3.47)
	Pooled PHEN/TPM: HR=1.02 and 95% CI = (0.30, 3.50)			

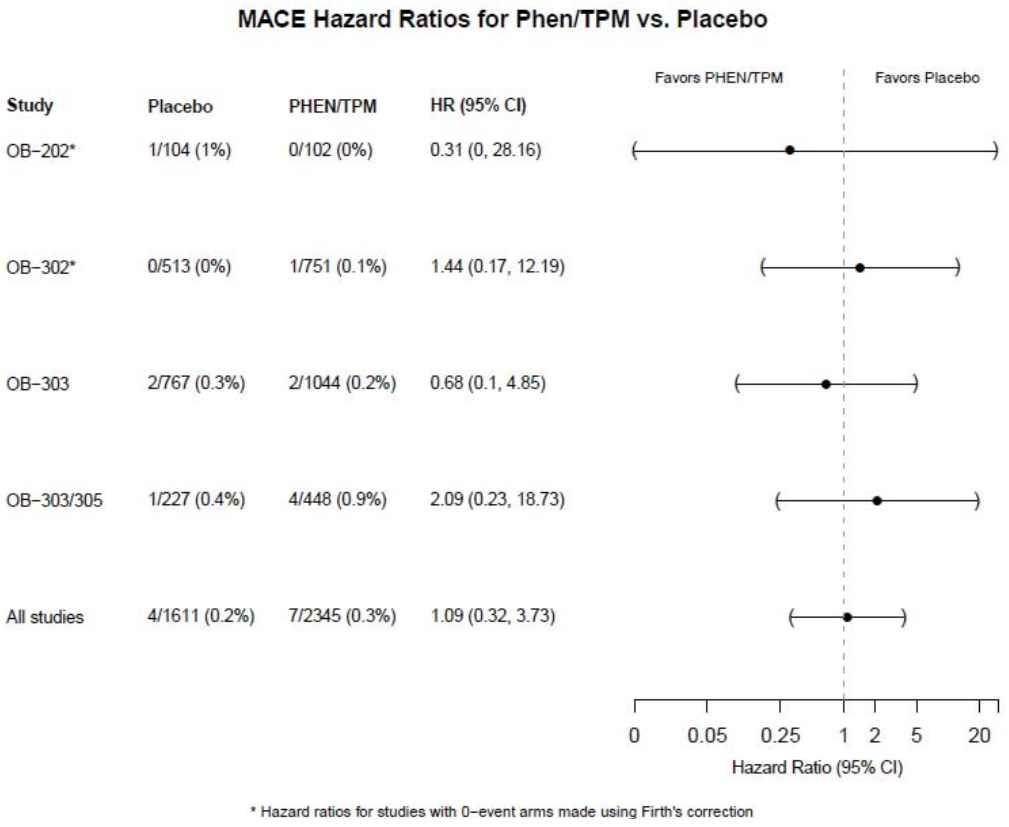
Finally, the unstratified analysis was repeated with the OB-305 data excluded due to the randomization concerns. The HRs and associated confidence intervals for this model are provided in Table 4.

Table 4: MACE Hazard Ratios vs. Placebo, No Stratification, OB-305 Data Excluded

	Placebo	**PHEN/TPM Low (3.75mg/23mg)**	**PHEN/TPM Mid (7.5mg/46mg)**	**PHEN/TPM High (15mg/92mg)**
N	1611	240	498	1607
MACE (%)	3 (0.18%)	1 (0.42%)	1 (0.20%)	3 (0.19%)
HR (95% CI)	-	2.05 (0.21, 19.72)	0.92 (0.10, 8.82)	0.92 (0.19, 4.55)
	Pooled PHEN/TPM: HR=1.03 and 95% CI = (0.25, 4.32)			

To summarize the results by study and the overall results, the HRs and associated confidence intervals were calculated for each study separately, with the doses pooled into a single treatment arm. The results are presented graphically in Figure 1.

Figure 1: Hazard Ratios and Confidence Intervals for MACEs by Study



There appears to be no significant evidence of a risk of MACEs from PHEN/TPM within one year of use, according to these analyses. This apparent lack of risk may extend up to two years of use, but due to the concerns regarding the randomization and selection bias in study OB-305, this result is not easily interpretable. When taken separately, the OB-303/305 study results appear to show a possible increase in MACE risk, but with the small sample size of subjects that participated in the OB-305 extension, this result is difficult to interpret. It is important to note that there are few MACEs in any of the arms, and that these trials were not designed or powered for MACE analyses, so interpretation of the MACE analysis results should be cautious.

3.2.2. Summaries of Heart Rate, Systolic Blood Pressure, Diastolic Blood Pressure, and Rate Pressure Product

The sponsor-provided cardiovascular risk analysis report gives descriptive statistics for the cardiovascular safety endpoints of mean change in heart rate, Systolic Blood Pressure (SBP), Diastolic Blood Pressure (DBP), and Rate Pressure Product (RPP)[1]. These statistics use Last Observation Carried Forward (LOCF) imputation; subjects that terminated the study early had their last recorded value carried forward to the end of the study period for the purpose of the estimated change. These tables of estimates are provided in Appendix I for ease of reference. There do not appear to be any significant increases in SBP, DBP, or RPP at one year. High-dose

[1] RPP is calculated as Systolic Blood Pressure times heart rate divided by 1,000.

PHEN/TPM appears to be related to a significant increase in mean change of heart rate at one year as compared to placebo: the difference in mean change of heart rate between high-dose PHEN/TPM and placebo is 1.60 bpm (95% CI: 0.88 to 2.32). A significant increase in heart rate was also found at the two-year endpoint for high-dose PHEN/TPM: the mean difference of 1.70 bpm (95% CI: 0.37 to 3.03); this was a statistically significant increase in heart rate, though not statistically different from that of the two-year placebo cohort (mean 0.4 bpm increase, 95% CI -0.88 to 1.67). Some increase in heart rate was shown in low- and mid-dose PHEN/TPM, as well, but these increases were not statistically significant; due to the low sample sizes in the low- and mid-dose arms, however, it is unclear whether this is evidence of no risk of heart rate increase or due to lack of power.

Assessment of the mean change at study-end, however, may not provide a full, accurate picture of cardiovascular safety; it is possible for possible safety effects to occur during the course of treatment that become less noteworthy by study-end and thus do not appear in these end-of-treatment summaries. Therefore, graphical summaries were produced for each of the four variables of heart rate, SBP, DBP, and RPP. With the assistance of clinical input, it was determined the most important cardiovascular safety population for these summaries was the Completers population: the population of subjects that completed the entire one- or two-year period on the randomized treatment. The Completers populations, then, are the ones represented in this longitudinal presentation of the data.

Figures 2 through 9 (provided in Appendix II) provide graphical summaries of the mean change in heart rate, Systolic Blood Pressure, Diastolic Blood Pressure, and Rate Pressure Product for the one-year Completers (studies OB-302 and OB-303) and two-year Completers (study OB-303/305). For ease of interpretation and because results did not appear to vary largely between treatment doses, graphical summaries are presented with the dosages pooled into a single treatment arm. There do not appear to be any notable differences between the active treatment and placebo other than a 1-2bpm mean increase in heart rate in the PHEN/TPM arm throughout the study, which supports the end-of-year difference discussed above. Also, there appear to be upward trends in SBP, DBP, and RPP in the second year for the OB-303/305 two-year Completers that may slightly mitigate the beneficial decreases in these vitals that appear at the end of the first year.

Overall, other than the possible safety signal of an increase in mean heart rate at end of treatment for high-dose PHEN/TPM users, there do not appear to be any significant concerns regarding increases in SBP, DBP, or RPP. It is worth noting, however, that subjects who completed two years of treatment did appear to see some increase in these endpoints between Weeks 56 and 108; that is, these endpoints tended to be somewhat higher at Week 108 than Week 56, though still overall reduced from baseline.

5. SUMMARY AND CONCLUSIONS

5.1 Conclusions and Recommendations

An analysis of the data from studies OB-202, OB-230, OB-302, OB-303, and OB-305 showed no notable differences in the hazard ratio of MACEs between any dose of PHEN/TPM and placebo, a result which was robust to numerous different models. Again, it should be noted that these studies were not designed or powered for MACE analyses, and so had few MACEs; the resulting analyses should be interpreted with caution due to a lack of power.

Descriptive analyses, including graphical representations, of systolic and diastolic blood pressure and rate pressure product for studies OB-302 and OB-303/305 did not appear to show any safety risks; each of these safety endpoints showed a mean decrease from baseline (indicating possibly improved cardiovascular health) at both one year and two years for the Completers population, and these decreases appeared to be similar to or larger than those seen in placebo subjects. It is worth noting, however, that at Week 108, these endpoints were higher than at Week 56 for two-year Completers, though still less than baseline. This may indicate that if there are desirable effects on cardiovascular health, they may become mitigated after more than one year of treatment.

A significant difference in change of heart rate from baseline after one year was found in the high-dose arm (mean increase of 1.6 bpm with standard error of 0.26) above the placebo (mean change of 0.0 bmp with standard error of 0.26). A similar increase was shown for the two-year data.

APPENDIX I

Table 5: Sponsor-provided summary of cardiovascular health markers for one-year cohort

Table 1. Changes in Blood Pressure (mmHg) and Heart Rate (bpm) From Baseline to Endpoint – Safety Set – 1-Year Cohort

Parameter Statistic	Placebo	PHEN/TPM 3.75/23	PHEN/TPM 7.5/46	PHEN/TPM 15/92
Systolic blood pressure (mmHg)				
n [1]	1532	234	488	1553
Baseline [2] mean (SD)	126.5 (13.25)	122.5 (11.11)	128.5 (13.63)	125.7 (13.12)
Endpoint [3] mean (SD)	124.3 (13.64)	119.1 (12.24)	123.4 (14.08)	120.5 (13.50)
Mean change (SD)	-2.1 (14.01)	-3.3 (11.95)	-5.2 (14.77)	-5.2 (14.48)
Within-treatment p-value [4]	<0.0001	0.0003	<0.0001	<0.0001
Comparison to placebo p-value [4]		0.2322	<0.0001	<0.0001
Diastolic blood pressure (mmHg)				
n [1]	1532	234	488	1553
Baseline [2] mean (SD)	79.6 (8.95)	77.8 (7.49)	80.6 (8.71)	79.0 (8.76)
Endpoint [3] mean (SD)	77.7 (9.62)	76.9 (8.24)	77.3 (8.82)	76.1 (8.82)
Mean change (SD)	-1.9 (9.61)	-0.9 (8.29)	-3.3 (9.87)	-2.9 (9.40)
Within-treatment p-value [4]	<0.0001	0.1402	<0.0001	<0.0001
Comparison to placebo p-value [4]		0.1362	0.0044	0.0023
Heart rate (bpm)				
n [1]	1532	234	488	1553
Baseline [2] mean (SD)	72.5 (9.58)	72.3 (9.22)	72.2 (10.07)	72.7 (9.87)
Endpoint [3] mean (SD)	72.5 (10.05)	73.6 (9.73)	72.7 (10.34)	74.3 (9.83)
Mean change (SD)	0.0 (10.19)	1.3 (10.32)	0.6 (10.18)	1.6 (10.28)
Within-treatment p-value [4]	0.9861	0.0499	0.2238	<0.0001
Comparison to placebo p-value [4]		0.0688	0.2933	<0.0001

Data from studies OB-202/DM-230, OB-302, and OB-303 are included.

1. n is the number of subjects with baseline and endpoint measurements.
2. Baseline is the last measurement obtained on or before the first dose date of double-blind study drug.
3. Endpoint is the last available measurement obtained during the double-blind treatment period.
4. P-values obtained from analysis of variance model with treatment as a fixed effect.

PHEN/TPM = VI-0521 fixed-dose combination of phentermine and topiramate; SD = standard deviation.

Source: Supporting Post-text Table 1

Table 6: Sponsor-provided summary of cardiovascular health markers for two-year cohort

Table 3. Changes in Blood Pressure (mmHg) and Heart Rate (bpm) From Baseline to Week 108 – Safety Set – 2-Year Cohort

Parameter Statistic	Placebo (N=227)	PHEN/TPM 7.5/46 (N=153)	PHEN/TPM 15/92 (N=295)
Systolic blood pressure (mmHg)			
n [1]	197	129	248
Baseline [2] mean (SD)	128.4 (14.45)	127.9 (11.71)	126.4 (13.61)
Week 108 mean (SD)	124.1 (12.44)	122.9 (13.44)	122.6 (12.81)
Mean change (SD)	-4.2 (15.12)	-5.0 (14.29)	-3.9 (14.00)
Within-treatment p-value [3]	<0.0001	<0.0001	<0.0001
Comparison to placebo p-value [3]		0. 6276	0.7760
Diastolic blood pressure (mmHg)			
n [1]	197	129	248
Baseline [2] mean (SD)	79.7 (9.55)	79.8 (9.09)	79.5 (8.69)
Week 108 mean (SD)	76.1 (9.34)	76.3 (8.76)	76.6 (8.24)
Mean change (SD)	-3.6 (10.27)	-3.5 (9.62)	-2.9 (9.44)
Within-treatment p-value [3]	<0.0001	<0.0001	<0.0001
Comparison to placebo p-value [3]		0.9477	0.4861
Heart rate (bpm)			
n [1]	197	129	248
Baseline [2] mean (SD)	70.6 (10.25)	72.2 (9.67)	73.0 (10.27)
Week 108 mean (SD)	71.0 (9.63)	73.4 (9.68)	74.8 (9.49)
Mean change (SD)	0.4 (9.86)	1.3 (10.17)	1.7 (10.64)
Within-treatment p-value [3]	0.5650	0.1654	0.0076
Comparison to placebo p-value [3]		0.4734	0.1771

1. n is the number of subjects with measurements at both baseline and Week 108.
2. Baseline is the last measurement obtained on or before the first dose date of double-blind study drug in study OB-303.
3. P-values obtained from analysis of variance model with treatment as a fixed effect.

PHEN/TPM = VI-0521 fixed-dose combination of phentermine and topiramate; SD = standard deviation.
Source: Supporting Post-text Table 2

Table 7: Sponsor-provided summary of Rate Pressure Product for one-year cohort

Table 9. Change in Rate-Pressure Product at Week 56 With LOCF – 1-Year Cohort, Heart Rate Subgroups, and Hypertension Subgroups

Treatment	N [1]	Baseline [2] Mean (SD)	Week 56 With LOCF [3] Mean (SD)	Change [4]			
				Mean (SD)	LS Mean (SE)	P-value	
						Within	vs Placebo
1-Year Cohort							
Placebo	1531	9.16 (1.55)	9.01 (1.61)	-0.15 (1.67)	-0.13 (0.05)	0.0173	
PHEN/TPM 3.75/23	234	8.86 (1.46)	8.77 (1.47)	-0.09 (1.54)	-0.19 (0.11)	0.0751	0.5470
PHEN/TPM 7.5/46	488	9.27 (1.59)	8.97 (1.63)	-0.30 (1.73)	-0.23 (0.08)	0.0044	0.1686
PHEN/TPM 15/92	1551	9.14 (1.58)	8.95 (1.52)	-0.19 (1.69)	-0.18 (0.05)	0.0007	0.3306
Subjects With Heart Rate Elevations							
Placebo	284	8.77 (1.70)	9.57 (1.84)	0.80 (1.70)	0.83 (0.11)	<0.0001	
PHEN/TPM 3.75/23	65	8.36 (1.46)	9.14 (1.71)	0.78 (1.57)	0.69 (0.21)	0.0010	0.5376
PHEN/TPM 7.5/46	132	8.99 (1.77)	9.42 (1.68)	0.44 (1.94)	0.54 (0.16)	0.0006	0.0798
PHEN/TPM 15/92	488	8.77 (1.73)	9.29 (1.68)	0.52 (1.69)	0.55 (0.09)	<0.0001	0.0118
Subjects Without Heart Rate Elevations							
Placebo	1247	9.25 (1.50)	8.88 (1.52)	-0.37 (1.59)	-0.42 (0.06)	<0.0001	
PHEN/TPM 3.75/23	169	9.06 (1.41)	8.63 (1.35)	-0.43 (1.39)	-0.55 (0.12)	<0.0001	0.2580
PHEN/TPM 7.5/46	356	9.38 (1.50)	8.80 (1.58)	-0.58 (1.56)	-0.58 (0.09)	<0.0001	0.0561
PHEN/TPM 15/92	1063	9.30 (1.48)	8.79 (1.42)	-0.51 (1.59)	-0.53 (0.06)	<0.0001	0.0478
Subjects With Hypertension at Baseline							
Placebo	616	9.66 (1.65)	9.28 (1.69)	-0.39 (1.79)	-0.34 (0.08)	<0.0001	
PHEN/TPM 3.75/23	33	9.76 (1.99)	9.22 (1.53)	-0.54 (1.76)	-0.39 (0.28)	0.1701	0.8597
PHEN/TPM 7.5/46	256	9.56 (1.64)	9.19 (1.68)	-0.37 (1.78)	-0.38 (0.12)	0.0017	0.6816
PHEN/TPM 15/92	633	9.56 (1.74)	9.09 (1.62)	-0.47 (1.85)	-0.49 (0.08)	<0.0001	0.0796
Subjects Without Hypertension at Baseline							
Placebo	915	8.83 (1.37)	8.83 (1.52)	0.00 (1.57)	0.02 (0.08)	0.8089	
PHEN/TPM 3.75/23	201	8.72 (1.30)	8.70 (1.45)	-0.02 (1.49)	-0.02 (0.12)	0.8644	0.7202
PHEN/TPM 7.5/46	232	8.96 (1.47)	8.73 (1.54)	-0.23 (1.66)	-0.18 (0.12)	0.1330	0.0633
PHEN/TPM 15/92	918	8.84 (1.39)	8.85 (1.44)	0.01 (1.53)	0.04 (0.08)	0.6481	0.7950

1. N is the number of subjects with values at both time points.
2. Baseline is the last measurement obtained on or before the first dose date of double-blind study drug.
3. Week 56 with LOCF is the last available measurement during the double-blind treatment period.
4. Least-squares mean, SE, and two-sided p-value are from an analysis of covariance model with treatment and study as fixed effects and baseline as a covariate.

LOCF = last observation carried forward; LS = least squares; PHEN/TPM = VI-0521 fixed-dose combination; SD = standard deviation; SE = standard error.

Sources: Supporting Post-text Tables 32, 33, 34, 35, and 36

APPENDIX II

Figure 2: Mean change in Heart Rate for OB-302 and OB-303 Completers

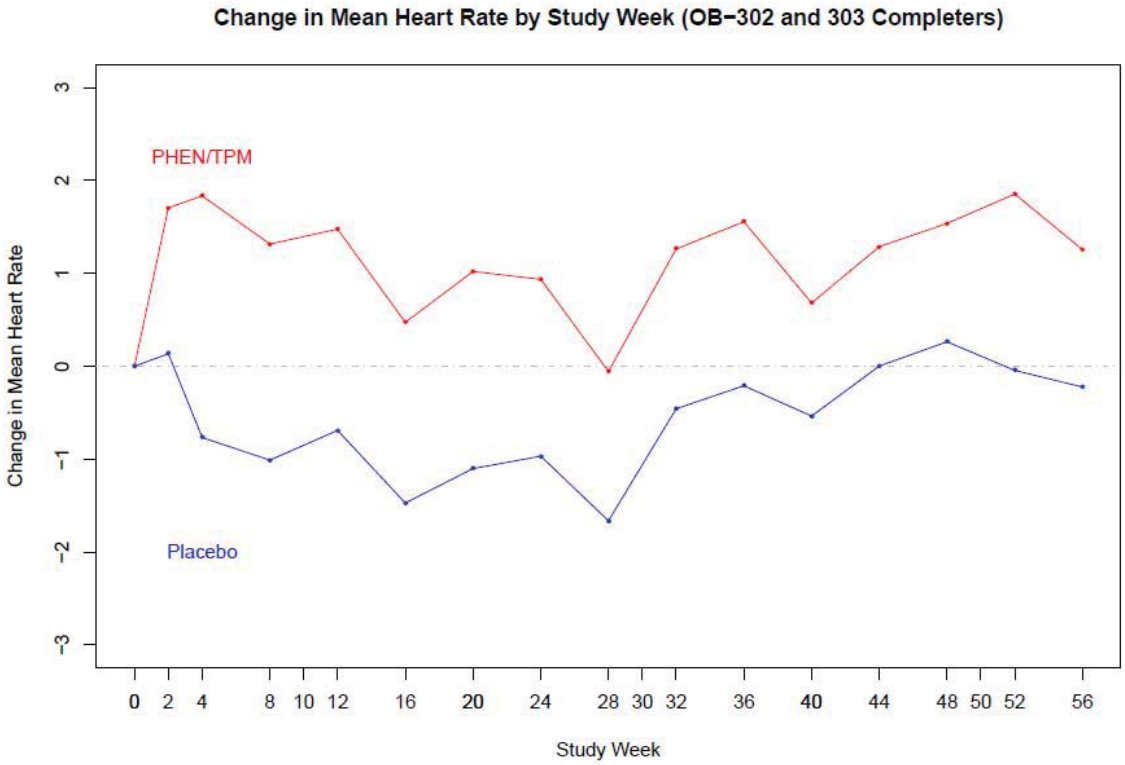


Figure 3: Mean change in Systolic Blood Pressure for OB-302 and OB-303 Completers

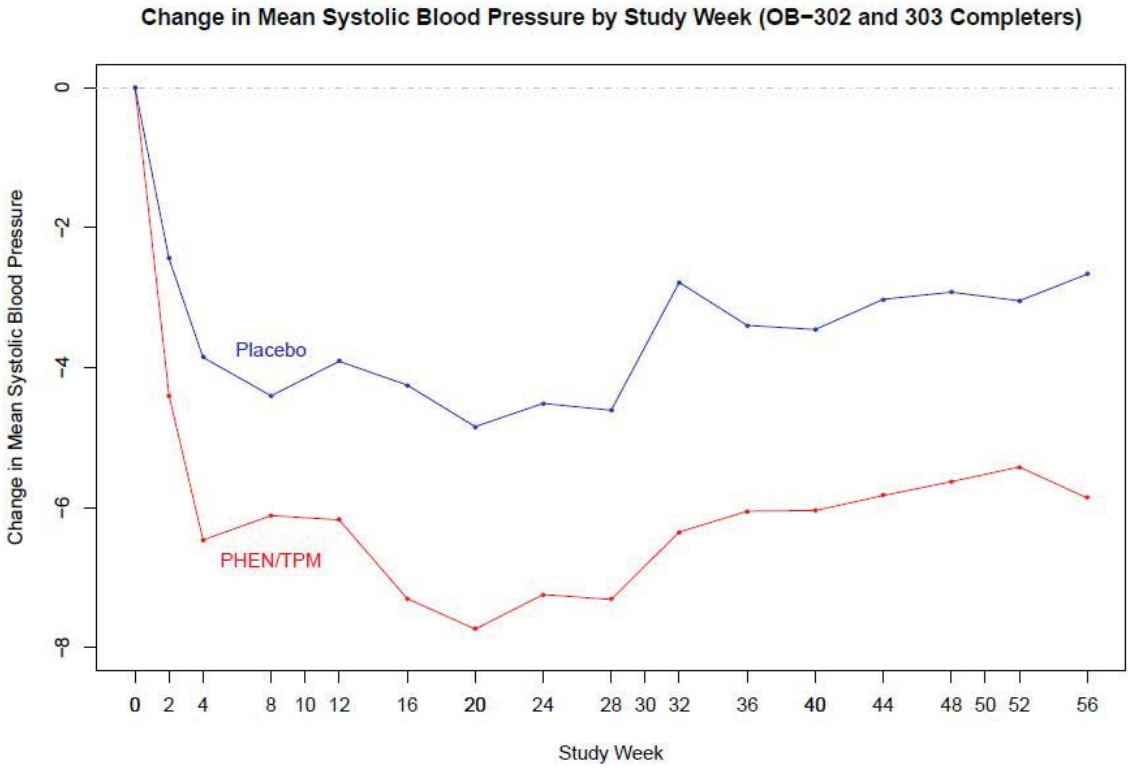


Figure 4: Mean change in Diastolic Blood Pressure for OB-302 and OB-303 Completers

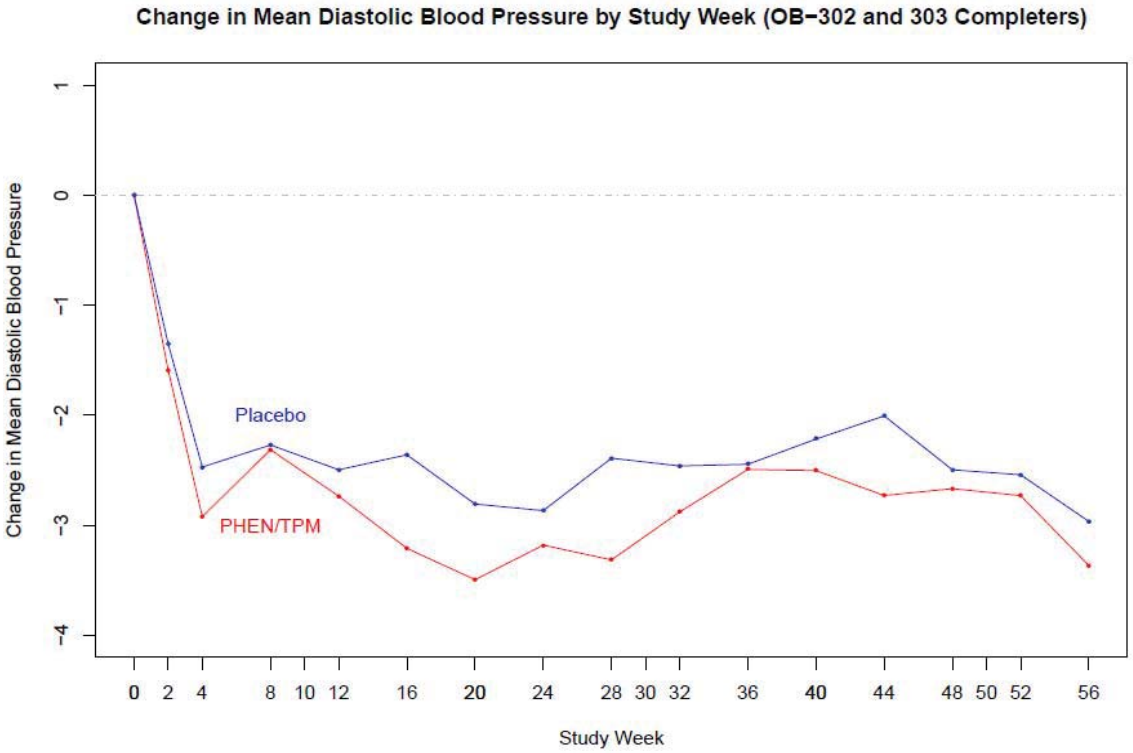


Figure 5: Mean change in Rate Pressure Product for OB-302 and OB-303 Completers

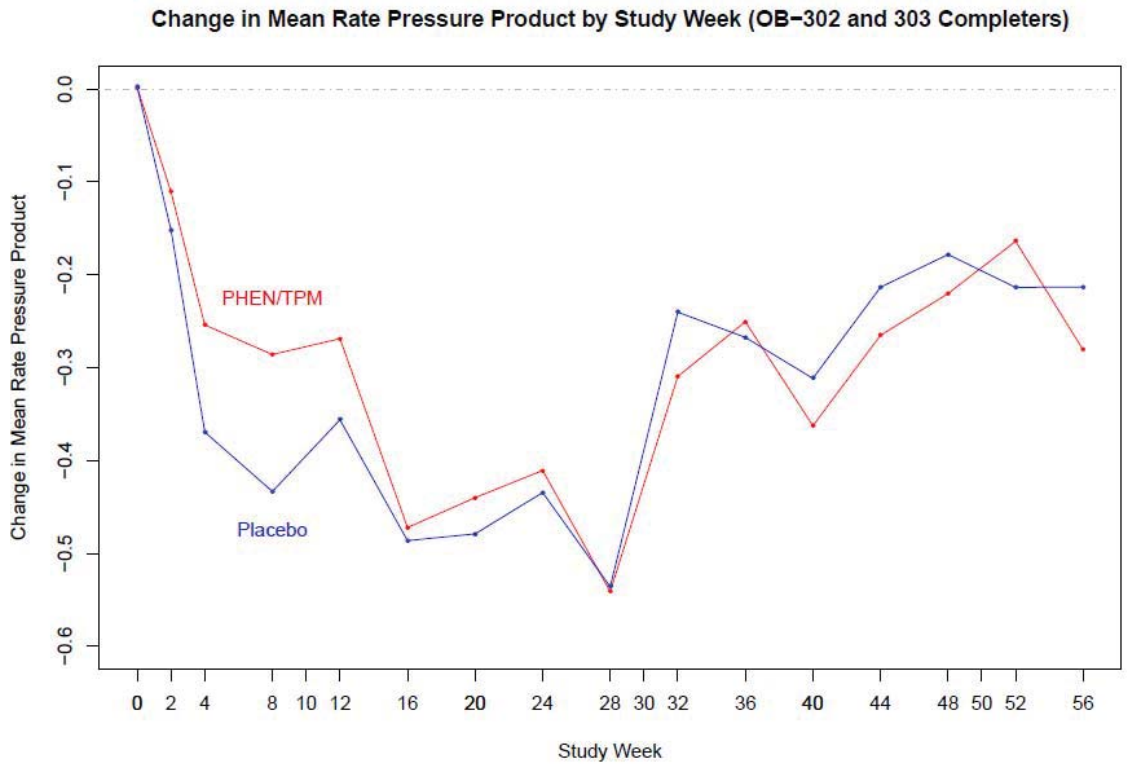


Figure 6: Mean change in Heart Rate for OB-303/305 Completers

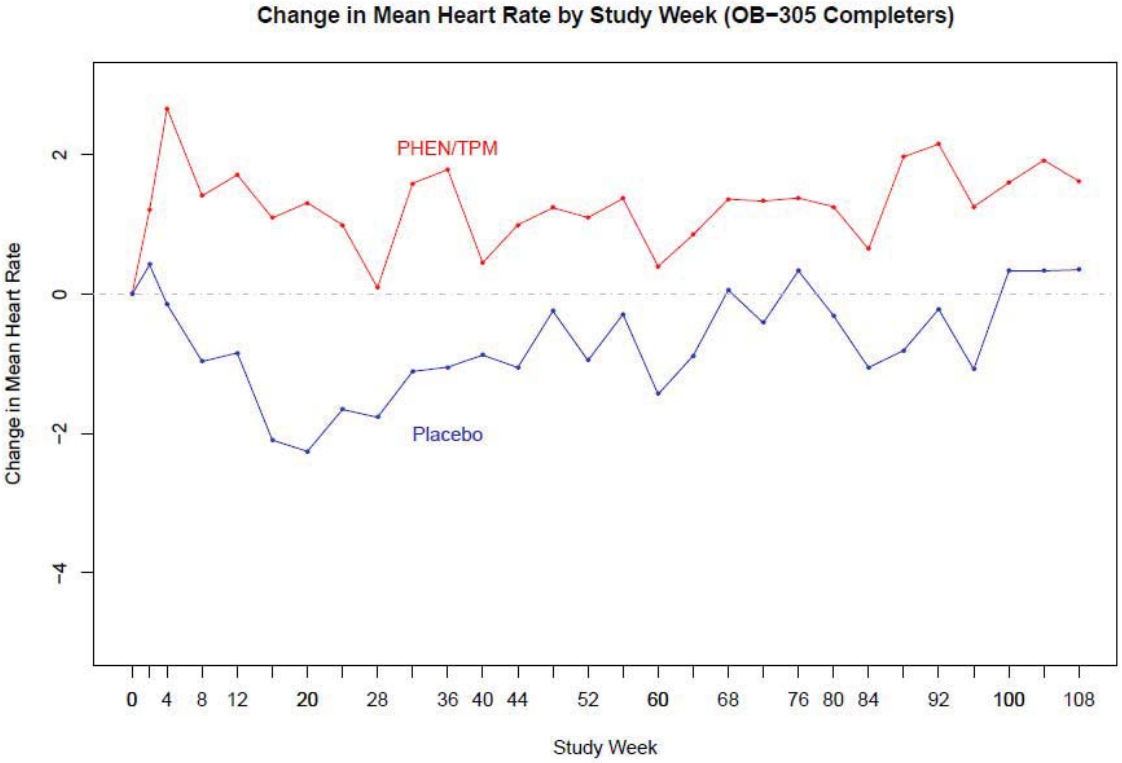


Figure 7: Mean change in Systolic Blood Pressure for OB-303/305 Completers

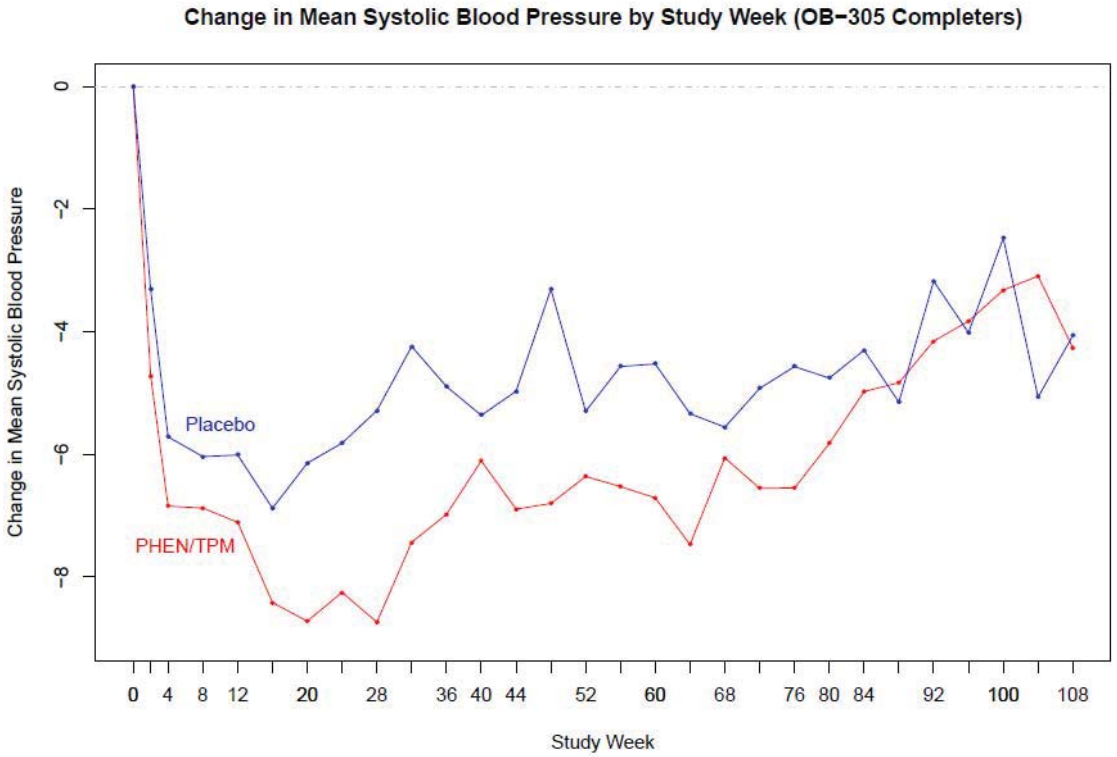


Figure 8: Mean change in Diastolic Blood Pressure for OB-303/305 Completers

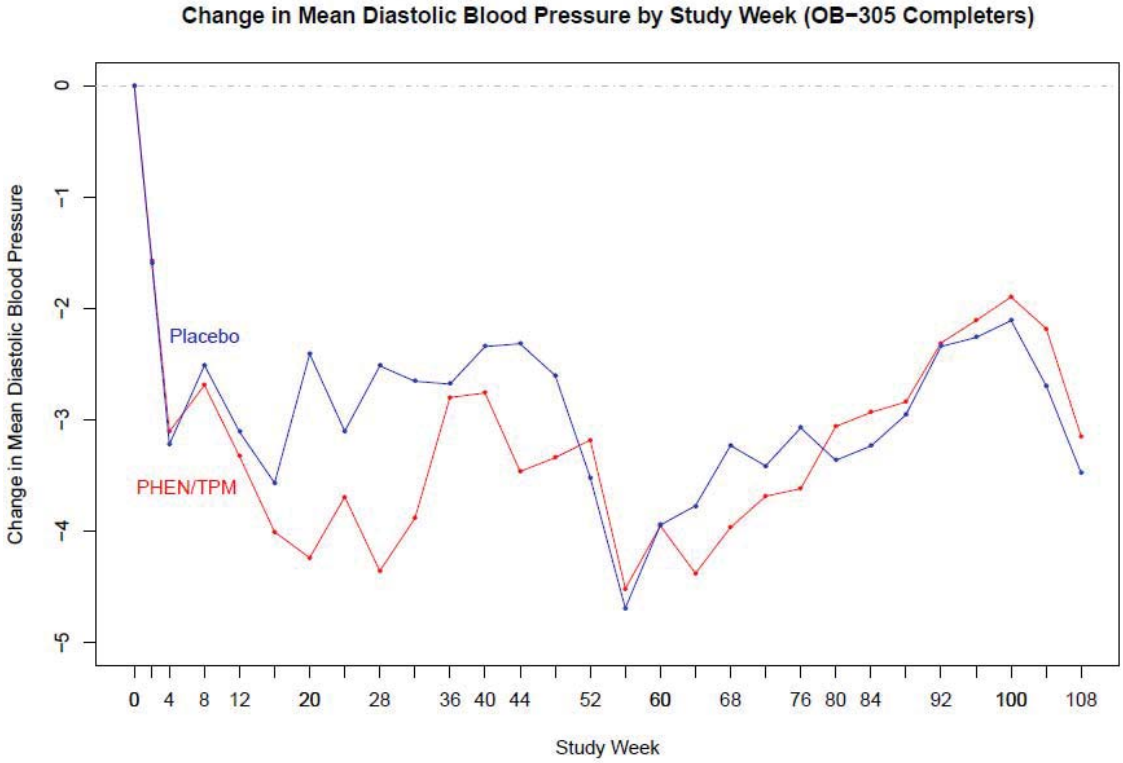
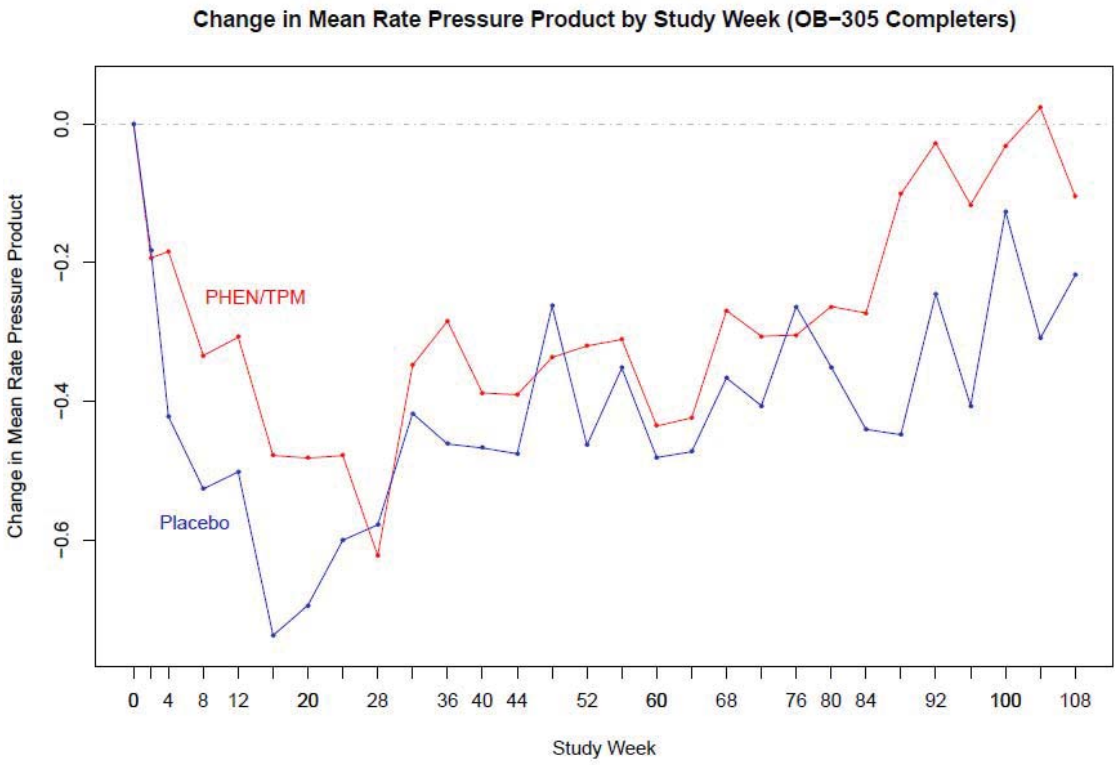


Figure 9: Mean change in Rate Pressure Product for OB-303/305 Completers



SIGNATURES/DISTRIBUTION LIST

Primary Statistical Reviewer: Benjamin Neustifter
Date: 3/20/12

Concurring Reviewer(s): Mat Soukup

Statistical Team Leader: Mat Soukup

Biometrics Deputy Division Director: Aloka Chakravarty

cc:
HFD-510/Dharia
HFD-510/Colman
HFD-510/Roberts
HFD-750/Neustifter
HFD-750/Soukup
HFD-750/Chakravarty
HFD-700/Patrician

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

/s/

BENJAMIN NEUSTIFTER
03/21/2012

ALOKA G CHAKRAVARTY
03/22/2012

MATTHEW J SOUKUP
03/22/2012
Concur with review

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

MEMORANDUM

NDA/Serial Number:	NDA 22-580/SN000
Drug Name:	Qnexa (phentermine and topiramate)
Indication(s):	Treatment of obesity in patients with and without weight-related comorbidities
Applicant:	Vivus, Inc.
Date Review Completed:	10/14/11
Statistical Reviewer:	Ben Neustifter, Ph.D., Biometrics Reviewer, DB VII
Concurring Reviewers:	Mat Soukup, Ph.D., Team Lead, DB VII
Medical Division:	Division of Metabolism and Endocrinology
Clinical Team:	Pooja Dharia, Regulatory Project Manager, DMEP Eric Colman, M.D., Clinical Team Leader, DMEP Mary Roberts, M.D., Clinical Reviewer, DMEP
Epidemiology Team:	Julia Ju, Pharm.D., Ph.D., Epidemiology Reviewer, DEPI I Diane Wysowski, Ph.D., Team Lead, DEPI I Judy Staffa, Ph.D., Division Director, DEPI II
Subject:	Statistical Review of Revised Protocol and Pooled Analysis Plan for OB-901, Fetal Outcomes Retrospective Topiramate Exposure Study (FORTRESS)

1. INTRODUCTION AND BACKGROUND

Qnexa is a fixed-dose combination of immediate-release phentermine (PHEN) and modified-release topiramate (TPM), submitted by the sponsor, Vivus Inc., indicated for weight management in obese (BMI ≥ 30 kg/m^2) and overweight (BMI 27-29.9 kg/m^2) individuals when accompanied by at least one weight-related comorbidity. Phentermine (PHEN) was approved in 1959 for the treatment of obesity, and has been indicated for short-term use only since 1973. Topiramate (TPM) was approved for the treatment of seizures in 1996 and for the prevention of migraine headache in 2004.

On May 25, 2011, the sponsor submitted Protocol OB-901 for the Fetal Outcomes Retrospective Topiramate Exposure Study (FORTRESS). Protocol OB-901 plans for a retrospective study of mother/infant dyads in administrative health care databases to estimate the risk of oral clefts (OCs) and major congenital malformations (MCMs) in infants where the mother was exposed to TPM in the first trimester. On July 19, 2011, the Division of Metabolism and Endocrinology (DMEP) submitted a response to the sponsor, including comments and suggestions on possible improvements to the statistical analysis portion of the protocol.

On September 22, 2011, the sponsor submitted a revised Protocol OB-901, including a separate draft document detailing the Summary Pooled Analysis Plan for FORTRESS. While this revised protocol, sponsor response, and Summary Pooled Analysis Plan address some of the concerns raised by the FDA, there remain portions of the proposed protocol and analysis that are of concern from a statistical standpoint.

2. Protocol OB-901

Much of Protocol OB-901 remains unchanged from the May 25, 2011 submission. Furthermore, some amendments to the protocol relate to areas outside the scope of this review and will not be discussed. Except for the noted portions below, the study objectives, methods, and plan remain unchanged from the May 25 submission. For reference on these aspects of FORTRESS, refer to the July 19, 2011, statistical review.

2.1 Study Objectives

These remain unchanged from the May 25, 2011, protocol submission:

- To estimate the Relative Risk (RR) of OCs in newborns of women exposed to TPM during the first trimester of pregnancy when compared to two control groups:

    - Newborns of women with remote prior exposure to TPM and other anti-epileptic drugs (AEDs), and
    - Newborns of women who were not exposed to TPM but had medical profiles similar to those who received TPM in the first trimester
- To estimate the RR of MCMs in newborns of women exposed to TPM during the first trimester of pregnancy when compared to the two above control groups

The secondary objectives for the study are:

- To estimate the risk of OCs and other MCMs in the newborns of women exposed to specific doses of TPM during the first trimester and, further, evaluate any dose response between TPM and the risk of OCs
- To estimate the RR of OCs and other MCMs in newborns of women exposed to TPM monotherapy as compared to women exposed to AED polytherapy regimens containing TPM
- To monitor for any signals of specific MCMs other than OC associated with TPM exposure in the first trimester
- To compare the proportion of infants born to mothers exposed to TPM who are born with low birth weight (LBW) as compared to
    - The proportion of infants born with LBW born to mothers not exposed to TPM
    - The proportion of all US newborns born with LBW
- To estimate the background risk of OCs and other MCMs for all mother-infant pairs from the administrative health care databases who are eligible for the study, and
- To estimate the positive predictive value (PPV) for automated claims diagnoses of OCs and MCMs overall

2.2 Study Plan

Much of the study plan remains the same from the May 25, 2011 submission. Data sources, however, have been altered to the following list:

- HealthCore Integrated Research Database
- OptumInsight's Normative Health Information Database
- Kaiser Permanente Research Databases (Northern and Southern California)
- Thomson Reuters MarketScan Medicaid Research Databases

Reviewer Note: By regaining HealthCore as a data source for the study (see Reviewer Note in July 19, 2011 statistical review), the sponsor has increased the expected sample size to 2,300 exposed mother-infant dyads. Section 2.4

will provide power calculations for comparisons to both control groups and both endpoints given this new sample size.

2.2.2 Exposure and Control Cohorts

The patient cohorts remain unchanged from the May 25, 2011 protocol submission. In brief resummarization, the cohorts are:

- **Exposure Cohort**: Mother-infant pairs exposed to TPM during the first trimester of pregnancy
- **Prior Exposure Cohort**: Mother-infant pairs for whom the mother was exposed to TPM or any other AED before the pregnancy of interest but not during the pregnancy
- **Diagnosis Control Cohort**: Mother-infant pairs with medical conditions similar to those in the Exposure Cohort but without exposure to TPM during the pregnancy. Dyads in this cohort may also be in the Prior Exposure Control Cohort.

2.3 Endpoints

The co-primary endpoints of Oral Clefts (OCs) and Major Congenital Malformations (MCMs) remain unchanged from the May 25, 2011 submission. As suggested in the comments by the FDA in the July 19, 2011 response, the sponsor's more restrictive alternative definitions of the endpoints are stated to be secondary. Other secondary endpoints remain the same from the May 25, 2011 protocol submission, and are acknowledged by the sponsor to be exploratory.

2.4 Study Size and Estimation of Study Power

According to the revised Protocol OB-901, the sponsor anticipates an estimated 2,300 exposed mother-infant pairs, with 16,1000 matched mother-infant pairs in the Diagnosis Control Cohort (7:1 ratio of control to exposed) and 10,000 mother-infant pairs in the Prior Exposure Cohort (approx. 4.8:1 ratio of control to exposed). The power calculations provided in the protocol are largely correct for the detection of differences in the Relative Risk (RR) of the endpoints in the Exposure Cohort compared to the Diagnosis Control Cohort. The protocol, however, does not provide the RR of the endpoints in the Exposure Cohort compared to the Prior Exposure Cohort.

Reviewer Notes: Based on protocol OB-901, the below power calculations assume that OCs have a background (not exposed to TPM during the first trimester) rate of 0.12%, or 1.2 events per 1,000 patients, and a background rate of 2.5%, or 25 events per 1,000 patients, for MCMs. They further assume a sample size of 2,300 mother-infant pairs in the Exposure Cohort, 16,100

dyads in the Diagnosis Control Cohort, and 10,000 in the Prior Exposure Cohort.

The following power calculations should be considered "best case" power scenarios for detecting difference between the Exposure Cohort and the control cohorts. Due to missing data or other problems, the amount of usable exposed dyads may be less than the number available in the data sources. In addition, the inflation of sample size using a 7-to-1 matching scheme in the Diagnosis Control Cohort will increase events but does not increase power at the same rate as a proportional increase in the exposed dyads. Variability between data sources can also reduce power. Thus, the power may be less than reported in these tables.

For many safety issues, a non-inferiority type of testing scheme is determined to be the appropriate approach. In the calculations that follow, it is assumed that the relative risk of the event of interest is 1 implying the risk is equivalent in the cohorts. Table 1 provides estimates of the smallest possible RR that could be ruled out under the study size and power restrictions, and the associated number of excess events above the background rate this represents for the two control cohorts.

Table 1: Comparison of Exposure Cohort to Control Cohorts, Assumed Relative Risk of 1

Control Cohort	Power	Oral Clefts		MCMs	
		Rule Out RR of	Excess Events*	Rule Out RR of	Excess Events*
Diagnosis	80%	3.30	2.8	1.30	7.5
	90%	3.97	3.6	1.35	8.8
Prior Exposure	80%	4.30	4.0	1.38	9.5
	90%	5.41	5.3	1.45	11.3

* Excess events in number of events per 1,000 patients above the background rate of 1.2 events per 1,000 (for OCs) or 25 events per 1,000 (for MCMs)

It can also be important to understand what the minimum detectable increase in relative risk may be if TPM exposure during the first trimester does cause an increase in OC/MCM risk as compared to the control cohorts. Table 2 assumes that the RR for the Exposure Cohort versus the Diagnosis Control Cohort is greater than 1 and provides the minimum detectable RR given the study size and power constraints.

Table 2: Comparison of Exposure Cohort to Control Cohorts, Elevated Risk for Topiramate

Control Cohort	Power	Oral Clefts		MCMs	
		Can Detect RR of	Excess Events*	Can Detect RR of	Excess Events*
Diagnosis	80%	2.94	2.3	1.30	7.5
	90%	3.41	2.9	1.35	8.75
Prior Exposure	80%	3.40	2.9	1.37	9.25
	90%	3.99	3.6	1.43	10.75

* Excess events in number of events per 1,000 patients above the background rate of 1.2 events per 1,000 (for OCs) or 25 events per 1,000 (for MCMs)

2.5 Statistical Analysis

As in the protocol dated May 25, 2011, statistical analysis in Study OB-901 will have two phases: Phase I will be the automated collection and analysis of data, and Phase II analyses will be repeated with an endpoint of cases where a chart review adjudicated "probable" or "possible" OCs.

Portions of the statistical analysis plan remain unchanged from the May 25, 2011 protocol, and will not be re-summarized.

2.5.1 Descriptive Statistics

The calculation of descriptive statistics remains unchanged from the May 25, 2011 protocol.

2.5.2 Analysis of Co-Primary Endpoints

Since all data come from healthcare data sources with restrictions upon the level and type of information that can be shared externally, all statistical analyses for the two co-primary endpoints (OCs and MCMs) will have two components: a preliminary within-site analysis and then a pooled final analysis conducted at a main site (RTI Health Solutions). For the preliminary analyses, each site will create tables of endpoints by exposure and other cross-classification variables. These tables will be created in two manners: by performing a "stratified analysis" by various covariates; and by using propensity scores to stratify subjects within each site.

2.5.2.1 Stratified Analysis

The first form of preliminary in-site analysis of data is the creation of cross-classification tables of exposure and study endpoint, using the potential confounders, including: maternal age, apparent indication, maternal diabetes mellitus, and concomitant exposure to AEDs and other potential teratogens.

Due to the low event rate of OCs, tables may only be by one or two stratification variables for each given table. The RRs for each stratum will be calculated as well as an overall RR of exposure versus control. The sponsor will attempt to identify confounding variables by defining covariates as confounding if the effect estimate for exposure is changed by at least 10% between the unstratified RR for exposure versus control and the stratified RRs. Any variables identified as confounders will be used for cross-classification tables to be sent for the pooled analysis.

Reviewer Comment: Due to the low event rate for OCs, the stratified method of identifying confounding variables may be unstable and/or uninformative for many variables. With low event counts, the shift of a single count from one cell to another may drastically change effect estimates and make the determination of whether a variable is identified as a confounder. Since this instability may mean that slightly different data would result in a very different final model and conclusion, this method of confounder identification for the final analysis may not be ideal. The sponsor should consider a propensity score analysis with quartile stratification as a preferable design, as well as a sensitivity analysis of propensity score matching (see following section). Furthermore, the use of 10% change in effect estimate as the threshold for determining whether a variable is a confounder should be justified if this analysis is to be included.

2.5.2.2 Propensity Score Analysis

The second form of preliminary table stratification proposed by the sponsor uses the calculation of propensity scores. The propensity scores will be calculated using a logistic regression model, where the odds of being exposed to TPM during the first trimester will be predicted using the following variables: Maternal age; infant sex; calendar year (using an indicator variable, not as a continuous variable); geographic region; tobacco smoking; dispensing or prescriptions of valproate, carbamazepine, phenytoin, other antiepileptic drugs, folic acid antagonists, or other known or suspected pharmaceutical teratogens (categories D or X); and history of claims or other mention of seizures, epilepsy, migraine, affective disorder, diabetes, hypertension, or obesity. Once the model has been fitted and a propensity score estimated for each subject, the data will then be adjusted as follows:

- The first percentile of the propensity scores among the exposed subjects will be made a cutoff point, and all subjects in both exposed and unexposed categories with propensity scores lower than this value will be excluded.
- The 99th percentile of the propensity scores among the unexposed subjects will also be made a cutoff and all subjects in both categories that have propensity scores larger than this value will be excluded.
- Counts for each exclusion will be recorded.

- For the remaining subjects, the data will stratified into tables by exposure cohort (unexposed, TPM exposure ≤ 100mg, TPM exposure > 100mg), study outcome, and propensity score (broken into deciles of the exposed distribution).

The sponsor notes that deciles may need to be merged in the final step if event counts are too sparse.

Reviewer Comment: Propensity scores are an effective method of adjusting for confounding in observational data, particularly when event counts are low. Due to the low event rate for OCs, however, the stratification by deciles in the final table will likely have a large number of 0-count cells, and prespecification of decile-merging may be difficult. The primary analysis is recommended instead to be based on propensity scores stratified into quartiles, to ensure sufficient cell and margin counts.

In addition to this change, a suggested sensitivity analysis would be to match subjects in-site by propensity scores, and have tables provided for pooled analysis of event counts across matched pairs (or many-to-one matching due to the larger control cohorts) stratified by exposure cohort. This would avoid the instability of the stratified analysis method. By only providing tables of counts of events by matched-pairs, the data sources should not be exceeding the bounds of restrictions upon information that may be given to external sources.

Multivariate logistic regression on the occurrence of MCMs may be an acceptable method for the analysis of the MCM endpoint. Covariates for this model should be pre-specified based on a clinical concern of their contribution to the event and kept to a small number to ensure adequate power. Propensity score analysis as described above may also be a useful analysis for the MCM data.

2.5.2.3 Pooled Analysis

The pooled data analysis will be performed at RTI Health Solutions using the stratified data provided from each individual site. Two pooled analyses are planned, using the stratified analysis site data and the propensity scores site data. The analysis using the stratified analysis site data will stratify by all the cofounders indicated by the 10% change rule described in Section 2.5.2.1 and use a Mantel-Haenszel-type estimator to determine if the RR varies significantly between exposure cohorts and/or sites. The sponsor notes that there may be difficulty with this method if too many strata have zero marginal counts. The analysis plan further states that precision may also be possibly improved by combining some strata of continuous variables with neighboring strata, which would be carried out by "examining shifts in the Mantel-Haenszel pooled estimator after trial versions of collapsing neighboring data." No

further information is given on pre-specification for combining neighboring strata with low/zero counts.

The second pooled analysis will be performed using the propensity score count tables detailed in section 2.5.2.2. As in the stratified analysis, a Mantel-Haenszel-type estimator will be used on a table stratified by propensity score deciles, site, exposure cohort, and study outcome. The pooled analysis protocol suggests that this analysis may also require merging of neighboring propensity score strata in order to obtain reasonable margin and cell counts.

Reviewer Comment: With the low event rate for OCs, it is reasonable to expect that there will be a non-trivial number of zero-count cells/margins in the cross-classification tables formed by the proposed stratified and propensity score analysis methods used in conjunction with the pooled final analysis. As stated in the comments to Section 2.5.2.2, a method that may be preferable would be to have the primary analysis be propensity scores stratified into quartiles, with a sensitivity analysis of in-site matching between subjects based on propensity scores and stratified tables based off these matched pairs. Creating tables with zero-count cells and margins and then attempting to combine neighboring strata may be difficult to accomplish without adding bias, and can be troublesome to pre-specify in a thorough manner. Therefore, the alternative method of using propensity scores stratified into fewer categories and a sensitivity analysis of matched subjects may be preferable.

2.5.3 Secondary Analyses

The secondary analyses remain unchanged from the May 25, 2011 protocol submission.

2.5.4 Sensitivity Analyses

The summary pooled analysis plan details several forms of sensitivity analysis that are planned:

- Alternative Definitions for Timing of Exposure, by changing the definition of when exposure during the first trimester may begin and end
- More Restrictive Definitions for OCs, based on claims and chart review, requiring both a diagnosis and a corrective procedure to be defined as OC
- More Restrictive Definitions for MCMs, that require both a diagnosis and corrective procedure to be defined as an MCM
- Impact of MCM Misclassification, by varying the sensitivity of the automated case definition for MCM and examining the RR
- Impact of Unmeasured Cofounding, by estimating the RR after the influence of an unmeasured confounder has been introduced using a previously designed spreadsheet program

Reviewer Comment: The sensitivity analyses seem acceptable from a statistical perspective. As they are exploratory, it will be difficult to extrapolate any conclusions from their results, but seeing results consistent with the main analyses may support the stability of those results.

It is recommended that a sensitivity analysis of matched subjects based on propensity scores be added to the study to support the propensity score stratification method.

3 COMMENTS THAT MAY BE CONVEYED TO THE SPONSOR

Comments on the revised protocol OB-901 and accompanying pooled analysis plan are made with the understanding that any unresolved issues regarding pre-specification of analysis or study design should be resolved *prior* to data collection and analysis.

Sample Size and Power Comments: With the re-introduction of the HealthCore database, the sample size and accompanying power are more favorable than the 1,400 estimated in the July 19, 2011 statistical review. Whether they are sufficient to proceed is a clinical decision.

Primary Analysis Comments: The following should be considered in revising your statistical methods for the primary analysis.

- The "stratified analysis" specified in the protocol may be quite unstable due to the low event rate for OCs. The shift of a single event from one cell to another may make noticeable differences in labeling a given variable as a "confounder" or not according to this method, leading to a possibly quite different conclusion drawn from the pooled analysis.
- If the stratified analysis is to be used despite this concern, the 10% threshold for determining a confounder should be well-justified.
- Propensity scores are an appropriate method of controlling for confounding when event rates are low. However, splitting the propensity score distribution into deciles for table analysis may result in many zero-count cells and margins. A more appropriate method would be to stratify the propensity scores into quartiles, rather than deciles. This should be supported by using propensity scores to perform subject matching within each site, then forming cross-classification tables of exposure and outcomes by subject pairs[1] as a sensitivity analysis.

[1] An example of this method can be found in the following publicly-available protocol: http://www mini-sentinel.org/work_products/Evaluations/AMI_Surveillance_Protocol_and_Appendices_ABC.pdf

- Multivariate logistic regression on the occurrence of MCMs would seem to be an acceptable method for the analysis of the MCM endpoint. Covariates for this model should be pre-specified based on a clinical concern of their contribution to the event and limited to ensure adequate power.

Benjamin Neustifter, Ph.D.
Statistical Reviewer, Biometrics 7

Concur: Mat Soukup, Ph.D.
Acting Team Leader, Biometrics 7

Cc:
NDA 22580
DB7/Chakravarty
DB7/Soukup
DB7/Neustifter
DMEP/Dharia
DMEP/Colman
DMEP/Roberts
DEPI I/Ju
DEPI I/Wysowski
DEPI II/Staffa

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

/s/

BENJAMIN B NEUSTIFTER
10/14/2011

MATTHEW J SOUKUP
10/14/2011
Concur with review

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

NDA/Serial Number:	NDA 22-580/SN000
Drug Name:	Qnexa (phentermine and topiramate)
Indication(s):	Treatment of obesity in patients with and without weight-related comorbidities
Applicant:	Vivus, Inc.
Date Review Completed:	
Statistical Reviewer:	Ben Neustifter, Ph.D., Biometrics Reviewer, DB VII
Concurring Reviewers:	Mat Soukup, Ph.D., Team Lead, DB VII
Medical Division:	Division of Metabolism and Endocrinology
Clinical Team:	Pooja Dharia, Regulatory Project Manager, DMEP Eric Colman, M.D., Clinical Team Leader, DMEP Mary Roberts, M.D., Clinical Reviewer, DMEP
Epidemiology Team:	Julia Ju, Pharm.D., Ph.D., Epidemiology Reviewer, DEPI I Diane Wysowski, Ph.D., Team Lead, DEPI I Judy Staffa, Ph.D., Division Director, DEPI II
Subject:	Statistical Review of Protocol for OB-901, Fetal Outcomes Retrospective Topiramate Exposure Study (FORTRESS)

1. INTRODUCTION AND BACKGROUND

Qnexa is a fixed-dose combination of immediate-release phentermine (PHEN) and modified-release topiramate (TPM), submitted by the sponsor, Vivus Inc., indicated for weight management in obese (BMI ≥ 30 kg/m^2) and overweight (BMI 27-29.9 kg/m^2) individuals when accompanied by at least one weight-related comorbidity. Phentermine (PHEN) was approved in 1959 for the treatment of obesity, and has been indicated for short-term use only since 1973. Topiramate (TPM) was approved for the treatment of seizures in 1996 and for the prevention of migraine headache in 2004.

On October 28, 2010, the Agency provided a Complete Response (CR) to the sponsor's application for Qnexa. The two reasons cited for the CR were: animal data and preliminary data from the North American Antiepileptic Drug Pregnancy Registry that suggest TPM poses teratogenic risk to women of child-bearing potential; and cardiovascular safety concerns based on a larger percentage of subjects in Phase 3 trials developing increases in heart rate in PHEN/TPM compared with placebo. Regarding the first concern, the actions for the sponsor listed by the CR letter included providing "a comprehensive assessment of topiramate's and phentermine/topiramate's teratogenic potential… [that] include[s] nonclinical and clinical data."

On May 25, 2011, the sponsor submitted Protocol OB-901 for the Fetal Outcomes Retrospective Topiramate Exposure Study (FORTRESS). Protocol OB-901 plans for a retrospective study of mother/infant dyads in administrative health care databases to estimate the risk of oral clefts (OCs) and major congenital malformations (MCMs) in infants where the mother was exposed to TPM in the first trimester.

2. Protocol OB-901

Protocol OB-901 for FORTRESS, dated May 25, 2011, provides an overview of the proposed study's methodologies and objectives, with some discussion of statistical analysis. While the specifics of the statistical analysis are not specified in Protocol OB-901 it is stated such detail will be forthcoming in an Analysis and Reporting Plan. Several points within the protocol are of statistical note and warrant further detail to reach agreement as described below.

2.1 Study Objectives

Protocol OB-901 lists two primary objectives and several secondary objectives for FORTRESS. The primary objectives are

- To estimate the Relative Risk (RR) of OCs in newborns of women exposed to TPM during the first trimester of pregnancy when compared to two control groups:
    - Newborns of women with remote prior exposure to TPM and other anti-epileptic drugs (AEDs), and
    - Newborns of women who were not exposed to TPM but had medical profiles similar to those who received TPM in the first trimester
- To estimate the RR of MCMs in newborns of women exposed to TPM during the first trimester of pregnancy when compared to the two above control groups

The secondary objectives for the study are

- To estimate the risk of OCs and other MCMs in the newborns of women exposed to specific doses of TPM during the first trimester and, further, evaluate any dose response between TPM and the risk of OCs
- To estimate the RR of OCs and other MCMs in newborns of women exposed to TPM monotherapy as compared to women exposed to AED polytherapy regimens containing TPM
- To monitor for any signals of specific MCMs other than OC associated with TPM exposure in the first trimester
- To compare the proportion of infants born to mothers exposed to TPM who are born with low birth weight (LBW) as compared to
    - The proportion of infants born with LBW born to mothers not exposed to TPM
    - The proportion of all US newborns born with LBW
- To estimate the background risk of OCs and other MCMs for all mother-infant pairs from the administrative health care databases who are eligible for the study, and
- To estimate the positive predictive value (PPV) for automated claims diagnoses of OCs and MCMs overall

Reviewer Note: While the protocol addresses all of the above objectives, the methods of analysis for the secondary objectives are brief and not strongly detailed. Due to this, and the status of the primary objectives as the endpoints of most clinical interest for regulatory action, the primary objectives will be the main focus of this review.

Reviewer Note: Protocol OB-901 uses the term "prevalence" to refer to "risk" (the incidence of an event divided by the sample size) throughout. This review will use risk as the appropriate statistical term.

2.2 Study Plan

OB-901 is designed to be a retrospective study with data taken from several large health care administrative databases. Data will be taken from the following data sources:

- Innovus' Ingenix Research Database
- Kaiser Permanente Research Database
- Thomson Reuters MarketScan Research Databases

Within each data source, data is proposed to be collected for mother/newborn dyads based on diagnosis codes and prescriptions in order to obtain both an exposure cohort and two control cohorts. OB-901 is proposed to have two phases, the first of which is based on automatic data collection, and the second involving review of selected medical records to confirm endpoints. The automated data collection (phase one) will allow for preliminary decisions on efficacy and safety to be made while the more lengthy medical review of records is carried out. The protocol thoroughly details how mothers' records will be linked to infants' records to provide accurate dyads.

Reviewer Note: Protocol OB-901 also lists the HealthCore Integrated Research Database (HIRD) as a data source, but a communication from the sponsor dated June 15, 2011, indicated that this database was no longer being used. The removal of HIRD from the set of data sources will have a possibly significant impact on the number of mother/infant dyads available for study, particularly in the cohort of dyads with TPM exposure during the first trimester. The impact of this reduction in size will be discussed in more detail in section 2.4.

2.2.1 Study Populations

The study population for OB-901 consists of women with a record of live birth during the period of 1997 through 2010 or 2011 (latest available data) and an identifiable newborn with at least 90-day post-delivery enrollment. The adult women must have 6 months of continuous enrollment in the health plan prior to the presumed conception date[1] and be between 15 and 49 years old on the delivery date. Exclusions will be made for mother-infant dyads with history of infection with one of the TORCH infections, alcohol abuse, substance abuse, or exposure to thalidomide or isotretinoin during the 6 months preceding estimated conception date or during pregnancy.

[1] The protocol includes an algorithm for estimating conception date based on diagnosis codes of premature birth or gestation length and provide a first trimester window based on this estimated date. Sensitivity analyses are included to determine how changes in this estimation method may affect results.

2.2.2 Exposure and Control Cohorts

Protocol OB-901 defines three patient cohorts for the study:

- **Exposure Cohort**: Mother-infant pairs exposed to TPM during the first trimester of pregnancy
- **Prior Exposure Cohort**: Mother-infant pairs for whom the mother was exposed to TPM or any other AED before the pregnancy of interest but not during the pregnancy[2]
- **Diagnosis Control Cohort**: Mother-infant pairs with medical conditions similar to those in the Exposure Cohort but without exposure to TPM during the pregnancy. Dyads in this cohort may also be in the Prior Exposure Control Cohort. The protocol states that 7 unexposed dyads will attempt to be matched to each pair in the Exposure Cohort. The matching criteria include apparent indication, maternal age (<35 or 35+), geographic region of health insurance plan, and calendar year of delivery.

Reviewer Comment: The choice to have two separate control cohorts in order to control for the effects of both prior TPM/AED exposure and indication is acceptable. While taking multiple matching pairs can lead to some increase in power, it is important to note that the exposed dyads are essential in the analysis and every attempt should be made to avoid discarding any exposed dyad. Thus, if some mother/infant exposed dyads do not have 7 available matched control pairs, it is recommended that the number of matched pairs for each exposed dyad be reduced, rather than eliminating any exposed pairs from the study.

Reviewer Note: Since the Diagnosis Control Cohort has several pairs matched to each Exposure Cohort pair, while the Prior Exposure Control Cohort does not, the Diagnosis Control Cohort will be several (seven as currently specified in the protocol) times larger than the Prior Exposure Control Cohort. Thus, there will be greater power to detect an increase in risk in the Exposure Cohort compared to the Diagnosis Control Cohort than to detect an increase in risk compared to the Prior Exposure Control Cohort.

2.3 Endpoints

The primary endpoint of OCs will be defined in two manners. A simple base definition of OC will be determined by ICD-9-CM diagnosis codes or a set of criteria related to Current Procedural Terminology (CPT) and procedure codes.

[2] This is operationalized by defining women in this cohort with an eligible dispensation of TPM or other AED at least 120 days before the estimated conception date, to ensure the prescription did not overlap with the first trimester.

Cases will be excluded if there is additional claims data that suggest syndromic malformations or genetic or chromosomal defects.

The second, more restrictive, manner of defining OCs will require both a diagnosis of oral cleft as well as a diagnosis code that indicates surgical repair of oral cleft within the first 365 days after birth. Both codes are considered to be equally relevant for purposes of evaluation.

The co-primary endpoint of MCMs will also have two definitions. Its more broad definition will use a set of ICD-9-CM diagnosis codes enclosed in the protocol. Exclusions will be made for mother/infant pairs with additional codes consistent with syndromic, genetic or chromosomal defects. As with OCs, a more restrictive definition will require an additional diagnosis code indicating a condition-specific procedure carried out within the first 365 days after birth.

Reviewer Note: The more restrictive definition of the primary endpoints will likely provide higher positive predictive value, but may exclude certain mother/infant dyads with less extreme cases of OC/MCM or ones that are corrected after the first year of the infant's life. As it is difficult to determine which definition of OC/MCM is most appropriate a priori, it is recommended that statistical analyses be performed with both definitions of OC/MCM.

The secondary endpoint of specific MCMs will be defined using the diagnosis codes for the specific MCMs of interest. The protocol states that no MCMs have been pre-specified for use as a secondary endpoint.

Low birth weight (LBW) will be defined as mother/infant pairs that contain diagnoses for "LBW" or "small for gestational age."

2.4 Study Size and Estimation of Study Power

Protocol OB-901, estimated that approximately 2,300 mother/infant dyads would be available to comprise the Exposed Cohort. Communication from the sponsor dated June 15, 2011, however, indicated that the sponsor had lost rights to the HealthCore database and that the Kaiser database contained approximately 100 fewer exposed pairs than expected. Thus, the most current estimate for the Exposed Cohort is 1,400 dyads.

Reviewer Note: The power calculations provided in OB-901 are based on an assumed Exposure Cohort of 2,300 patients and incorrect assumptions (in particular, the assumption regarding whether the RR for OCs/MCMs in topiramate is greater than 1, indicating a risk, is stated to be true when their calculations assume the RR equals 1), this review will not provide the power calculations from the protocol. Instead, power calculations will be calculated by the review under various assumptions. In particular, the power

calculations below are provided for an Exposure Cohort of 1,400 pairs as well as for an Exposure Cohort of 2,200. The latter Exposure Cohort size is applicable if the sponsor is able to regain rights to the HealthCore database for the study or replace the HealthCore database with other data sources.

Reviewer Comment: The sponsor provides calculations for the probability that the upper confidence bound for the RR will lie below certain benchmarks if the true RR is 1. The calculation method for these probabilities should be provided to the FDA for further review.

Reviewer Notes: Based on protocol OB-901, the below power calculations assume that OCs have a background (not exposed to TPM during the first trimester) rate of 0.12%, or 1.2 events per 1,000 patients, and a background rate of 2.5%, or 25 events per 1,000 patients, for MCMs.

Finally, these power calculations should be considered "best case" power scenarios for detecting difference between the Exposure Cohort and the Diagnosis Control Cohort. Due to missing data or other problems, the amount of usable exposed dyads may be less than the number available in the data sources. In addition, the inflation of sample size using a 7-to-1 matching scheme will increase events but does not increase power at the same rate as a proportional increase in the exposed dyads. Variability between data sources can also reduce power. Thus, the power may be less than reported in these tables.

For many safety issues, a non-inferiority testing scheme is determined to be the appropriate approach. In the calculations that follow, it is assumed that the relative risk of the event of interest is 1 implying the risk is equivalent in the cohorts. Table 1 provides estimates of the smallest possible RR that could be ruled out under the study size and power restrictions, and the associated number of excess events above the background rate this represents for the Diagnosis Control Cohort. Note that because the Diagnosis Control Cohort has a 7-to-1 matching scheme for the control cohort, these results do not apply to the Prior Exposure Control Cohort; tables for the latter cohort will be provided subsequently.

Table 1: Comparison of Exposure Cohort to Diagnosis Control Cohort, Assumed Relative Risk of 1

Exposed Dyads	Power	Oral Clefts		MCMs	
		Rule Out RR of	Excess Events*	Rule Out RR of	Excess Events*
1,400	80%	4.47	4.2	1.40	10.0
	90%	5.66	5.6	1.48	12.0
2,200	80%	3.40	2.9	1.31	7.9
	90%	4.12	3.7	1.37	9.3

* Excess events in number of events per 1,000 patients above the background rate of 1.2 events per 1,000 (for OCs) or 25 events per 1,000 (for MCMs)

It can also be important to understand what the minimum detectable increase in relative risk may be if TPM exposure during the first trimester does cause an increase in OC/MCM risk as compared to the Diagnosis Control Cohort. Table 2 assumes that the RR for the Exposure Cohort versus the Diagnosis Control Cohort is greater than 1 and provides the minimum detectable RR given the study size and power constraints.

Table 2: Comparison of Exposure Cohort to Diagnosis Control Cohort, Elevated Risk for Topiramate

Exposed Dyads	Power	Oral Clefts		MCMs	
		Can Detect RR of	Excess Events*	Can Detect RR of	Excess Events*
1,400	80%	3.75	3.3	1.39	9.8
	90%	4.43	4.1	1.46	11.5
2,200	80%	3.01	2.4	1.30	7.5
	90%	3.49	3.0	1.36	9.0

* Excess events in number of events per 1,000 patients above the background rate of 1.2 events per 1,000 (for OCs) or 25 events per 1,000 (for MCMs)

Tables 3 and 4 summarize the same information for a comparison between the Exposure Cohort and the Prior Exposure Control Cohort. Recall that the Prior Exposure Cohort will be of the same size as the Exposure Cohort. As such, the margin of risk that can be ruled with this comparative group is higher than the comparison of the Exposure Cohort to the Diagnosis Control Cohort.

Table 3: Comparison of Exposure Cohort to Prior Exposure Control Cohort Assumed Relative Risk of 1

Exposed Dyads	Power	Oral Clefts		MCMs	
		Rule Out RR of	Excess Events*	Rule Out RR of	Excess Events*
1,400	80%	21.26	24.31	1.95	23.75
	90%	34.36	40.03	2.17	29.25
2,200	80%	11.46	12.56	1.71	17.75
	90%	16.80	18.96	1.86	21.5

* Excess events in number of events per 1,000 patients above the background rate of 1.2 events per 1,000 (for OCs) or 25 events per 1,000 (for MCMs)

Table 4: Comparison of Exposure Cohort to Prior Exposure Control Cohort, Elevated Risk for Topiramate

Exposed Dyads	Power	Oral Clefts		MCMs	
		Can Detect RR of	Excess Events*	Can Detect RR of	Excess Events*
1,400	80%	5.56	5.5	1.77	19.25
	90%	6.50	6.6	1.91	22.75
2,200	80%	4.48	4.2	1.60	15
	90%	5.07	4.9	1.71	17.75

* Excess events in number of events per 1,000 patients above the background rate of 1.2 events per 1,000 (for OCs) or 25 events per 1,000 (for MCMs)

2.5 Statistical Analysis

Phase 1 of Study OB-901 (the automated collection and analysis of data) will contain analysis of all OCs identified by diagnosis code, as well as the more restrictive definition of OCs discussed in Section 2.3. In the second phase of the study, analysis will be repeated only for those cases where a chart review adjudicated "probable" or "possible" OC. MCMs, on the other hand, will only be analyzed using the automated data definition.

Reviewer Note: Given the likely small number of OCs, it is important to check for sensitivity of the results to changes in definition. Therefore, using both the base and restricted definitions in phase one may be useful, as discussed in section 2.3, until the results from medical adjudication are received. Because the greater number of MCMs will preclude adjudication of all such records, it may be acceptable to only use automated data for the MCM analyses.

2.5.1 Descriptive Statistics

Protocol OB-901 states that descriptive statistics will be calculated by study cohort for both demographic variables and possibly relevant covariates. Categorical variables will be summarized by frequencies and proportions, and continuous variables will be summarized by their minima, maxima, and 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles. The risk of OCs and other MCMs will also be reported for each cohort.

2.5.2 Analysis of Co-Primary Endpoints

As stated in Section 2.1, the primary objectives of Study OB-901 are to estimate the Relative Risk (RR) of OCs and MCMs in the Exposure Cohort as compared to the two Control Cohorts. Protocol OB-901 defines the main analysis of this outcome as a "stratified analysis." This procedure will be performed by stratifying event rates between cohorts with 1 to 2 other variables of interest (e.g. maternal age, apparent indication, maternal diabetes) and reporting the event rate in strata if they differ from the unstratified event rates by at least 10%. The sponsor states that in this manner, covariates that are confounded with the outcomes may be uncovered.

The protocol also states that propensity scores will be calculated by predicting TPM exposure within each cohort. Further information on this analysis is stated to be forthcoming in an Analysis and Reporting Plan; the protocol states only that the propensity scores will be used to "assess overlap between the treatment and the comparison cohort."

Since MCMs will likely have a larger event size, the analysis will also include multivariate regression to model the RR of MCMs. The covariates for such a model are not specified in the protocol.

Reviewer Comment: Sufficient details of the primary analysis have not been provided in the protocol and, as such, agreements cannot be made on the analysis at this time. As only a limited number of OC events are anticipated to be observed, we have concerns about the degrees of freedom available to perform any stratified analysis. Therefore, we believe that a propensity score analysis seems to be the most useful statistical analysis method for the small number of events anticipated in OCs. Propensity scores should be calculated by modeling the risk of TPM exposure during the first trimester based on covariates of possible influence (such as maternal age, indication, etc.). These propensity scores can then be used to match subjects from the Exposure Cohort to the two Control Cohorts, and then an analysis method such as McNemar's test can be used on these pairs to attempt to discern a difference in event rates between the cohorts.

Multivariate logistic regression on the occurrence of MCMs may be an acceptable method for the analysis of the MCM endpoint. Covariates for this model should be pre-specified based on a clinical concern of their contribution

to the event and kept to a small number to ensure adequate power. Propensity score analysis as described above may also be a useful analysis for the MCM data.

2.5.3 Secondary Analyses

The protocol for Study OB-901 lists several secondary analyses to be performed on the phase 1 (automated) data:

- The main endpoint analyses will be re-performed only for those women with indications for migraine
- A dose-response relationship between TPM and OCs/MCMs will be explored by dichotomizing TPM usage into average daily doses of more than 100mg and 100mg or less during the first trimester. No further information on this analysis is given
- A duration-response relationship between TPM and OCs/MCMs will also be explored by evaluating event rates compared to the number of TPM-exposed days during the first trimester
- The main analyses will be repeated to compare TPM monotherapy versus TPM as part of a multi-drug regimen

The second phase data will also be used to calculate the PPV of the diagnostic and procedure codes for OCs. The PPV of MCM codes will also be calculated using a restrictive automated definition of MCMs.

Reviewer Comment: The proposed secondary analyses appear to be subgroup comparisons which will be problematic due to the small number of events anticipated for OCs and MCMs. As such it is likely that the proposed secondary analyses will have little power to detect any actual signals for the relationships they purport to assess. If the intent is to utilize such analyses in an inferential manner, a multiplicity adjustment should be provided for the multiple tests being performed, otherwise the analyses should be listed as exploratory in nature.

2.5.4 Sensitivity Analyses

The protocol defines some sensitivity analyses to examine the robustness of the results to varying exposure and outcome definitions. The OC analyses will be repeated using only phase 2 (adjudicated medical records) cases that are rated as "probable" (thus eliminating "possible" cases that were used in the main analyses). Furthermore, the window for first trimester estimation (to determine first trimester TPM exposure) will be moved in order to assess the effect of its definition on results.

The protocol also states that sensitivity analyses for an unmeasured cofounder's effects on analyses, and for potential misclassification of MCMs will be carried out, but the specifics of these analyses are not given.

Reviewer Comment: Sensitivity analyses may be useful. It is difficult to comment further on the sensitivity analyses in the absence of a more specific plan for analysis.

2.5.5 Signal Detection

The protocol calls for an exploratory analysis of the risk of MCM by organ system affected. The analysis will compare the RR of organ-specific MCMs between the Exposure Cohort and an unspecified control group, with a signal defined as an RR of 5 or greater when there are at least 4 events across datasets.

Reviewer Comment: This analysis may be useful, but will likely have low power due to the stratification of MCMs by organ system. It is unlikely that any significant results will be obtained by this exploratory analysis. As an exploratory analysis, this assessment of MCM risk by organ system will have limited regulatory utility.

2.5.6 Birth weight analysis

Using the ICD-9 definitions of low birth weight (LBW) and small for gestational age, an exploratory analysis of LBW is planned. The risk of LBW will be reported across cohorts and RRs will be calculated between the Exposure and Control cohorts. A further, unspecified, analysis will be conducted that is focused on exposure during the third trimester (the time of most fetal weight gain). The proportion of LBW infants in the Exposure Cohort will also be compared to the national standard.

Note: As with the secondary and sensitivity analyses, it is difficult to comment without further information on the types of analyses. The protocol states that the national standard for LBW is 8%, so it is possible to have acceptable power with these analyses. It may be difficult to determine exposure during the third trimester based on medical records. As above, these will be viewed as exploratory analyses and will have limited regulatory utility.

3 COMMENTS THAT MAY BE CONVEYED TO THE SPONSOR

All comments on Protocol OB-901 are made with the understanding that the statistical concepts and analyses in this protocol are incomplete and should be fully specified in a Statistical Analysis Plan prior to data collection and analysis.

Cohort Comment: *The choice to have two separate control cohorts in order to control for the effects of both prior TPM/AED exposure and indication is*

acceptable. While taking multiple matching pairs can lead to some increase in power, it is important to note that the exposed dyads are essential in the analysis and every attempt should be made to avoid discarding any exposed dyad. Thus, if some mother/infant exposed dyads do not have 7 available matched control pairs, it is recommended that the number of matched pairs for each exposed dyad be reduced, rather than eliminating any exposed pairs from the study.

Sample Size and Power Comments: The sample size calculations provided in the protocol appear to have changed based upon your June 15, 2011 communication. The following should be considered in your determination of sample size.

- A clear assumption of the assumed relative risk should be provided along with appropriate terminology to define the objective being assessed (e.g. ruling out a given RR rate under non-inferiority, or detecting a minimum increase with a certain amount of power).
- Absent the HealthCore data base, there is a concern about the power of the study to rule out a clinically meaningful level of risk.

Due to the inability to reproduce your sample size calculations and the concern about the power of the study, you should submit a revised protocol addressing the points above in order to reach agreement on the sample size projections for this study.

Primary Analysis Comments: The following should be considered in revising your statistical methods for the primary analysis.

- The "stratified analysis" specified in the protocol would not allow for easily interpretable inference of the results. With such a small number of events anticipated for OCs, it will be common for the apparent relative risk in a stratum to shift based on a difference of very few events from one strata to another. Also, it will have little power to detect true signals. This analysis will likely be difficult to interpret.
- A propensity score analysis may be a useful statistical analysis method for the small number of events anticipated in OCs. Propensity scores could be calculated by modeling the risk of TPM exposure during the first trimester based on covariates of possible influence (such as maternal age, indication, etc.). These propensity scores can then be used to match subjects from the Exposure Cohort to the two Control Cohorts, and then an analysis method such as McNemar's test can be used on these pairs to attempt to discern a difference in event rates between the cohorts.
- Multivariate logistic regression on the occurrence of MCMs would seem to be an acceptable method for the analysis of the MCM endpoint. Covariates for this model should be pre-specified based on a clinical concern of their contribution to the event and limited to ensure adequate power. Propensity score analysis may also be useful.

Sensitivity/Secondary Analysis Comments:

- It is likely that the secondary analyses in the protocol will have little power to detect any actual signals for the relationships they are testing for in OCs, and so are recommended to be seen as exploratory rather than inferential. While the number of MCMs may be greater and so these tests may have more power for this endpoint, explicit power calculations may need to be performed to ensure their viability as analyses before treating them as more than exploratory.
- The signal detection and birth weight analyses will be seen as exploratory and will thus have limited regulatory utility.

Benjamin Neustifter, Ph.D.
Statistical Reviewer, Biometrics 7

Concur: Mat Soukup, Ph.D.
Acting Team Leader, Biometrics 7

Cc:
NDA 22580
DB7/Chakravarty
DB7/Soukup
DB7/Neustifter
DMEP/Dharia
DMEP/Colman
DMEP/Roberts
DEPI I/Ju
DEPI I/Wysowski
DEPI II/Staffa

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

/s/

BENJAMIN B NEUSTIFTER
07/19/2011

MATTHEW J SOUKUP
07/19/2011
Concur with review

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

Statistical Review and Evaluation

CARCINOGENICITY STUDIES

IND/NDA Number:	NDA 22-580
Drug Name:	Phentermine HC1
Applicant:	Sponosor: VIVUS, Inc. 1172 Castro Street Mountain View, California 94040
	Test Facility: (b) (4)
Documents Reviewed:	Electronic data submitted on March 24, 2010, Also include the sponsor's reports submitted.
Review Priority:	Standard
Biometrics Division:	Division of Biometrics -6
Statistical Reviewer:	Min Min, Ph.D.
Concurring Reviewer:	Karl Lin, Ph.D.
Medical Division:	The Division of Metabolism and Endocrinology Products
Reviewing Pharmacologist:	David Carlson, Ph.D.
Project Manager:	Pooja Dharia
Keywords:	Carcinogenicity, Dose response

Table of Contents

1. ..... Background 3

2. ..... Rat Study 3
2.1. Sponsor's analyses 4
2.1.1. Survival analysis 4
2.1.2. Tumor data analysis 6
2.2. Reviewer's analyses 7
2.2.1. Survival analysis 7
2.2.2. Tumor data analysis 7

3. ..... Summary 9

4. ..... Appendix 11

5. ..... References: 33

1. Background

In this submission the sponsor included reports of one animal carcinogenicity study in rats. This study was intended to assess the carcinogenic potential of the test article, phentermine HC1 in rats when administrated daily by gavage for about 104 weeks. Results of this review have been discussed with the reviewing pharmacologist Dr. Carlson.

2. Rat Study

Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and two control groups. Three hundred CD® [Crl:CD® (SD)] rats of each sex were randomly allocated to treated and control groups. Three treated groups (60 animals/sex/group) received the test article, phentermine HC1, daily by gavage for an intended period of 104 weeks. The intended dose levels were 3, 10 or 30 mg eq./kg/day. Two additional groups of 60 animals/sex/group served as the controls and received the vehicle, distilled water. The first control group was a vehicle control and the second control group was a pair-fed control receiving a fixed amount of food based on the food consumption of animals at 30 mg/kg/day. The test article or vehicle was administered to all groups via oral gavage, once a day for up to 105 consecutive weeks, at a dose volume of 10 mL/kg. The study design is as the following tables:

Group	Dose Level (mg/kg/day)	Volume (mL/kg)	Number of Animals	
			Males	Females
1	Control	10	60	60
2	3	10	60	60
3	10	10	60	60
4	30	10	60	60
5	Control (pair-fed)[a]	10	60	60

[a] Beginning in Week 3, animals in Group 5 will be pair fed according to the amount of food consumed by animals in Group 4.

Observations for morbidity, mortality, injury, and the availability of food and water were conducted twice daily for all animals. Beginning in Week 53, a third mortality check in the evening was conducted throughout the remainder of the study. Observations for clinical signs and masses were conducted weekly. Body weights were measured and recorded weekly for the first 26 weeks of the study and every two weeks thereafter. Food consumption was measured and recorded weekly. Serological health screens were conducted on five randomly selected animals per sex prior to study start. Blood for clinical pathology evaluations were collected from all surviving animals at the terminal necropsy and animals euthanized *in extremis*, when possible. Necropsy examinations were performed under procedures approved by a veterinary pathologist on animals euthanized *in extremis*, found dead or dying prior to euthanasia, and all surviving animals at the scheduled necropsy (Week 98 and Week 105 for females and males, respectively).

For fatal and incidental neoplasms, the onset date was considered to be the fate date of the affected animal. For mortality independent neoplasms, the onset date was considered to be the first appearance of a related abnormality (e.g., abrasion, nodule, and/or swelling) that was consistently recorded in the detailed clinical

observation data at the site of the neoplasm. The first appropriate mass observation was used as the onset date if no appropriate clinical observations were previously noted for a given neoplasm site. If neither an appropriate clinical observation nor a mass finding could be correlated to a given neoplasm, the onset date was considered to be the fate date of the affected animal.

2.1. Sponsor's analyses

2.1.1. Survival analysis

Male and female data were analysed separately. Intercurrent mortality data was analyzed using the Kaplan-Meier product-limit method. An overall test comparing all groups was conducted using a log-rank test. If this overall test was significant ($p < 0.05$) and there were more than two groups, then a follow up analysis was done where each treatment group was compared to the control group using a log-rank test. Results of all pair-wise comparisons are reported at the 0.05 and 0.01 significance levels. All endpoints were analyzed using two-tailed tests.

Sponsor's findings: The Kaplan-Meier product-limit survival curves from the sponsor's report are presented in Figure 1 and 1A for males and females, respectively. Survival at study termination for males (Week 105) was 41.7%, 38.3%, 58.3%, 53.3%, and 58.3% in the vehicle control, 3, 10, 30 mg/kg/day dosage levels and the pair-fed control group, respectively. Survival at study termination for females (Week 98) was 33.3%, 40.0%,
50.0%, 48.3%, and 48.3% in the vehicle control, 3, 10, 30 mg/kg/day dosage levels and the pair-fed control group, respectively. No difference in survival was seen between the high dose of 30 mg/kg/day and the pair-fed control group.

No test article-related effects were seen on survival during this 2-year study. Survival in the treated groups was comparable or higher than the vehicle control throughout the study. The survival in females in the vehicle control group was lower than all other female groups during the last 6 months of the study resulting in the early terminal sacrifice of all female groups in Week 98 when the total female survival in the vehicle control group reached 20 animals. All male groups survived to their scheduled termination after 104 weeks of the study.

Figure 1: Kaplan-Meier plot of Survival in Male Rats

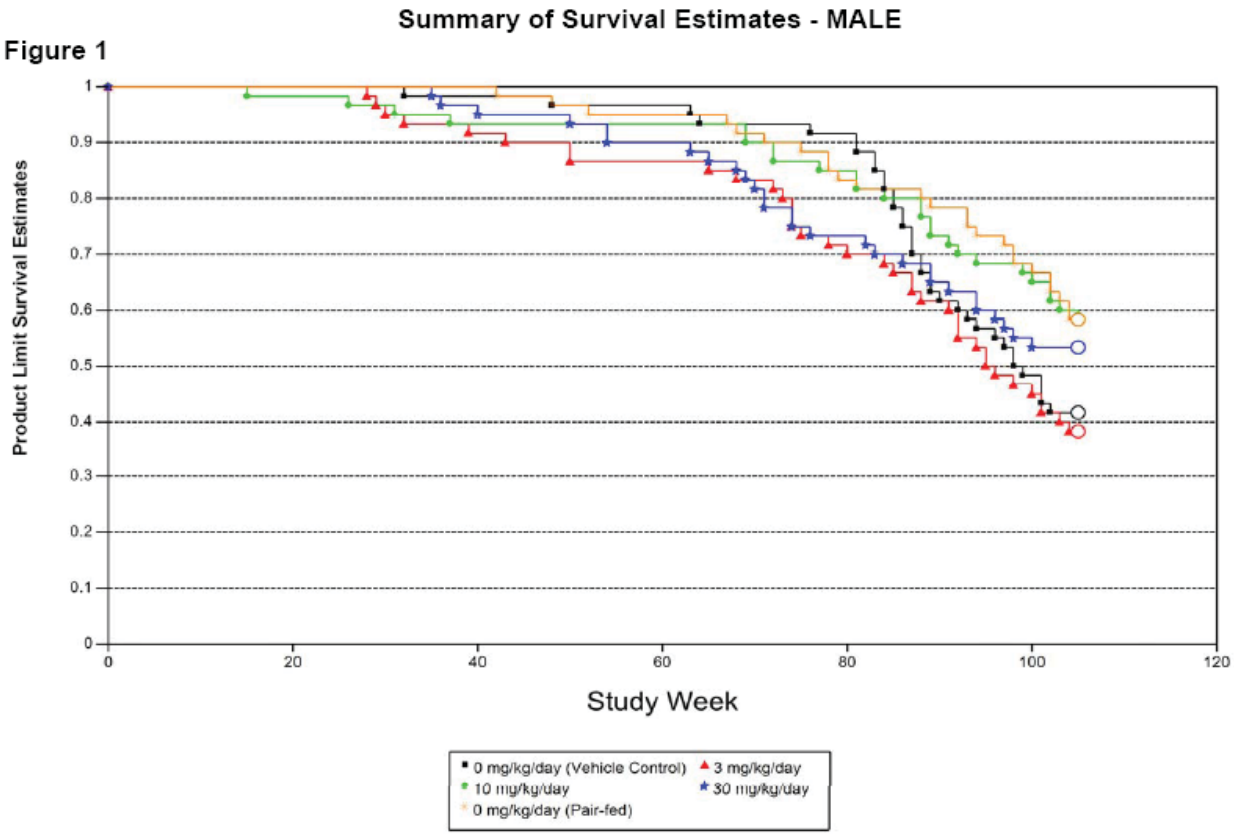
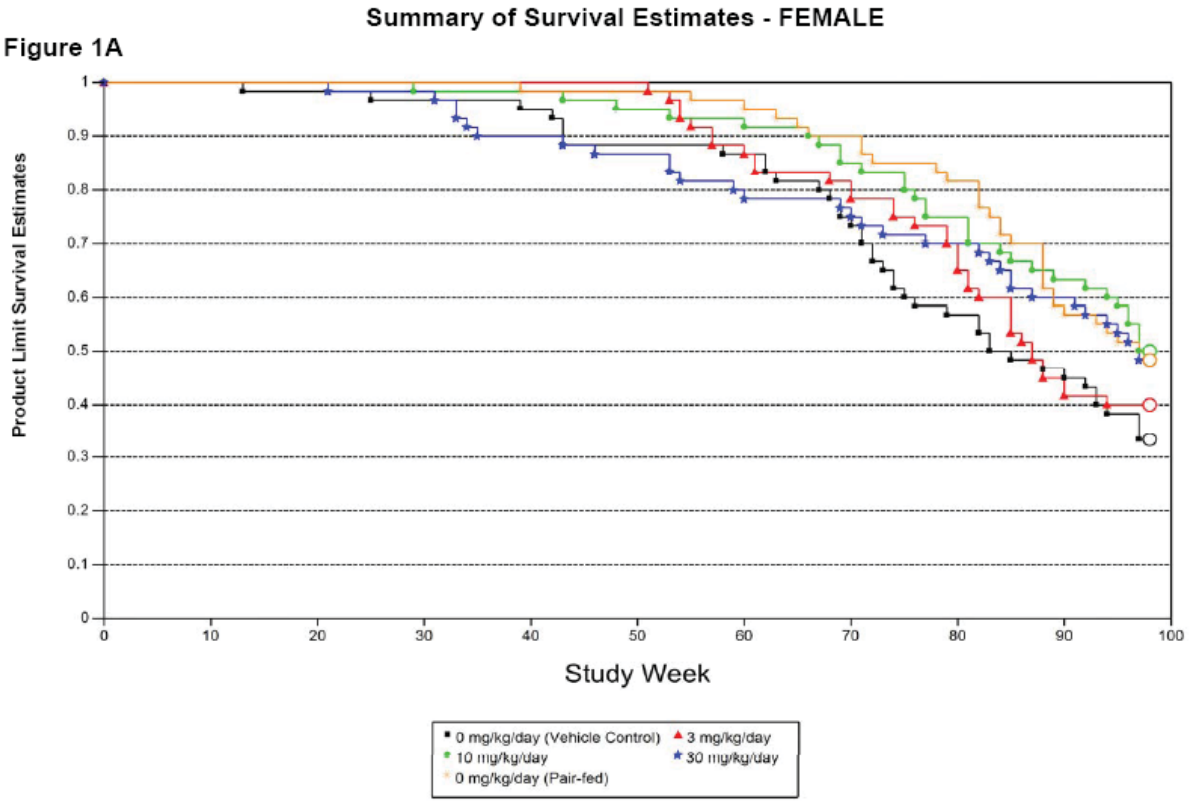


Figure 1A: Kaplan-Meier plot of Survival in Female Rats



2.1.2. Tumor data analysis

Tumor incidence data was analyzed using both survival adjusted and unadjusted tests. The unadjusted tests were based on the incidence and number of sites examined for each tumor type. The Cochran-Armitage trend test was calculated and Fisher's exact test was used to compare each treatment group with the control group. The survival adjusted test was conducted according to the prevalence/mortality methods described by Peto *et al.*

Sponsor's findings: There were no test article-related neoplastic microscopic observations noted. No statistically significant increase (with one exception detailed in next paragraph) was found by either trend tests or pair-wise comparisons by the defined significance levels, in the incidence of any tumor type in any phentermine hydrochloride treatment group when compared with incidence rates in the vehicle control and pair-fed control groups.

One exception was benign granular cell tumors (a common tumor) present in the uterus of 30 mg/kg/day female rats. The Cochran-Armitage trend test was statistically significant (p-value 0.0039) when compared to the vehicle control group; however, p-values were above significance levels by all other statistical analyses of uterine granular cell tumors. The incidences of benign granular cell tumors in the uterus were 0/60, 0/60, 3/60, 5/60, and 3/60 in vehicle control, 3 mg/kg/day, 10 mg/kg/day, 30 mg/kg/day, and pair-fed control

groups, respectively. In addition, the percent incidence range of benign granular cell tumors in the uterus in historical control groups from previous 2 year rat studies conducted at (b) (4) (Sprague Dawley Rat Neoplastic Data 2 Year Studies 09-01 Audited) was well within (0.0% to 10.0%) the observed percent incidence in the female 30 mg/kg/day group (8.3%). This finding was not considered to be test article related. In conclusion, the daily oral administration of phentermine HCl for 2 years to (b) (4) CD rats at dose levels of 3, 10, and 30 mg/kg/day did not produce any evidence of a carcinogenic effect.

2.2. Reviewer's analyses

To verify sponsor's analyses and to perform the additional analysis suggested by the reviewing pharmacologist, this reviewer independently performed survival and tumor data analyses. Data used in this reviewer's analyses were provided by the sponsor electronically. Three sets of analysis have been done only for the vehicle control.

2.2.1. Survival analysis

The survival distributions based on the animals in all four treatment groups (three treated groups and one vehicle control group or three treated groups and pair-fed control group) were estimated by the Kaplan-Meier product limit method. The dose response relationship and homogeneity of survival distributions were tested using the Cox test (Cox, 1972) and the Generalized Wilcoxon test (Gehan, 1965). The intercurrent mortality data are given in Tables 1A and 1B in the appendix for males and females, respectively. The Kaplan-Meier curves for survival rate are given in Figures 1A (1), 1A (2), 1B (1) and 1B (2) in the appendix for males and females, respectively. Results for the tests for dose response relationship and homogeneity of survivals, are given in Tables 2A and 2B in the appendix for males and females, respectively.

Reviewer's findings: The tests results showed no statistically significant dose-response mortality in both males and females when the treated groups were compared separately with the pair-fed control and vehicle control, respectively or across all treatment groups.

2.2.2. Tumor data analysis

The tumor data were analyzed for dose response relationships and pair-wise comparisons of the control group with each of the treated groups were performed using the Poly-k method described in the paper of Bailer and Portier (1988) and Bieler and Williams (1993). Two sets of analysis were performed. The first set used the vehicle control and the second set the pair-fed control in the tests for dose response and pairwise comparisons. One critical point for Poly-k test is the choice of the appropriate value of k. For long term 104 week standard rat and mouse studies, a value of k=3 is suggested in the literature. Hence, this reviewer used k=3 for the analysis of this data. For the calculation of p-values the exact permutation method was used. The tumor rates and the p-values of the tested tumor types are listed in Tables 3A (1), 3A (2), 3B (1) and 3B (2) in the appendix for males and females, respectively.

Multiple testing adjustment: Adjustment for the multiple dose response relationship testing was done using the criteria developed by Lin and Rahman (1998). The criteria recommend the use of a significance level α=0.025 for rare tumors and α=0.005 for common tumors for a submission with two species, and a significance level α=0.05 for rare tumors and α=0.01 for common tumors for a submission with only one species study in order to keep the false-positive rate at the nominal level of approximately 10%. A rare tumor is defined as one in which the spontaneous tumor rate is less than 1%. The adjustment for multiple pair-wise comparisons was done using the criteria developed by Haseman (1983) that recommends the use of a significance level α=0.05 for

rare tumors and α=0.01 for common tumors, in order to keep the false-positive rate at the nominal level of approximately 10%.

It should be noted that the recommended test levels by Lin and Rahman for the adjustment of multiple testing were originally based on the result of a simulation and an empirical study using the Peto method for dose response relationship analysis. However, some later simulation results by Rahman and Lin (2008) indicate that the criteria apply equally well to the analysis using the poly-3 test.

The following list of tumor combinations suggested by the reviewing pharmacologist were also tested by this reviewer:

hemangiomas and hemangiosarcomas from all sites (include separate analysis for hemangiomas and hemangiosarcomas)
mesotheliomas (all sites)
leukemias (all sites, male rats)
lymphomas (all sites)
chondroma / osteosarcoma / osteoma (all bone-- e.g., bone, cranium, femur)
lipoma / liposarcoma at same tissue site
kidney tubular cell adenomas / carcinomas (male rats)
liver hepatocellular adenomas / carcinomas (male rats)
pancreas islet cell adenoma and mixed acinar / islet cell adenoma (male rats)
pancreas mixed acinar / islet cell adenoma and acinar cell adenoma (male rats)
pituitary anterior lobe adenoma / carcinoma (male rats, female rats)
skin and subcutis basal cell adenoma / carcinoma (male rats)
skin and subcutis squamous cell papilloma / carcinoma / keratoacanthoma (male rats, female rats)
testis interstitial cell adenoma / mesothelioma / rete testis adenoma / sex cord stromal tumor (male rats)
thymus thymoma (begnin and malignant) (male rats)
thyroid c-cell adenoma / carcinoma
thyroid follicular cell adenoma / carcinoma
duodenum leiomyoma / leiomyosarcoma (female rats)
mammary adenoma / carcinoma (female rats)
mammary fibroadenoma / fibrocarcinoma (female rats)
uterus stromal polyp / sarcoma (female rats)
uterus adenoma / adenocarcinoma (female rats)
uterus/vagina stromal neoplasms
pituitary adenomas / carcinomas anterior lobes
oral cavity/tonggue squamous cell papillomas/ carcinomas

Reviewer's findings: Following tumor types showed p-values less than or equal to 0.05 either tests for dose response relationship and/or pair-wise comparisons between the vehicle control or pair-fed control and each of individual treated groups.

Tumor Types with P-Values ≤ 0.05 for Dose Response Relationship or Pair-wise Comparisons (Control, low, medium and high dose groups)

Organ Name	Tumor Name	control	Low	Med	High	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Vehicle control									
Female brain	ASTROCYTOMA, MALIGNA	1	0	1	4	0.028	1.000	0.803	0.237
uterus with cer	GRANULAR CELL TUMOR,	0	0	3	5	0.005	.	0.164	0.038
Male testes	ADENOMA, INTERSTITIA	5	2	2	8	0.039	0.921	0.948	0.234
Pair-fed control									
Female cavity, thoraci	HIBERNOMA,BENIGN+MAL	0	1	1	3	0.038	0.472	0.500	0.105
uterus with cer	GRANULAR CELL TUMOR,	3	0	3	5	0.050	1.000	0.651	0.297

Based on the criteria of adjustment for multiple testing of trends proposed by Lin and Rahman, the positive dose-response relationship in the incidence of granular cell tumor in uterus when vehicle control was used and of combination of benign and malignant hibernoma in thoracic cavity when pair-fed control was used in females were considered to be statistically significant. Also based on the criteria by Haseman, the increased tumor incidences of granular cell tumor in uterus in high dose group in female rats when compared to vehicle control group was considered to be statistically significant.

3. Summary

In this submission the sponsor included reports of one animal carcinogenicity study in rats. This study was intended to assess the carcinogenic potential of the test article, phentermine HC1 in rats when administrated daily by gavage for about 104 weeks.

Rat Study: Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and two control groups. Three hundred CD® [Crl:CD® (SD)] rats of each sex were randomly allocated to treated and control groups. Three treated groups (60 animals/sex/group) received the test article, phentermine HC1, daily by gavage for an intended period of 104 weeks. The intended dose levels were 3, 10 or 30 mg eq./kg/day. Two additional groups of 60 animals/sex/group served as the controls and received the vehicle, distilled water. The first control group was a vehicle control and the second control group was a pair-fed control receiving a fixed amount of food based on the food consumption of animals at 30 mg/kg/day. The test article or vehicle was administered to all groups via oral gavage, once a day for up to 105 consecutive weeks, at a dose volume of 10 mL/kg.

The tests showed no statistically significant dose-response mortality in both males and females when the treated groups were compared separately with pair-fed water and vehicle control groups, respectively, or across all treatment groups.

The tests showed statistically significant positive dose response relationship in incidence of granular cell tumor in uterus using vehicle control and of combination of benign and malignant hibernoma in thoracic cavity using

pair-fed control in females. Pair-wise comparisons showed statistically significantly increased incidences of granular cell tumor in uterus in high dose group in female rats when compared to the vehicle control group.

Min Min, Ph.D.
Mathematical Statistician

Concur: Karl Lin, Ph.D.
Team Leader, Biometrics-6

cc:
Archival NDA 22-580
Dr. David Carlson
Dr. Tiwari
Dr. Nevius
Dr. Machado
Dr. Lin
Dr. Min

4. Appendix

Table 1A: Intercurrent Mortality Rate
Male Rats

	VEHICLE CONTROL		LOW		MEDIUM		HIGH1		PAIR-FED CONTROL	
Week	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT
0-52	2	3.3%	7	11.7%	4	6.7%	4	6.7%	3	5.0%
53-78	3	8.3%	9	27.1%	5	15.0%	12	26.7%	6	15.0%
79-92	18	38.3%	7	39.0%	8	28.3%	6	36.7%	4	21.7%
93-104	12	58.3%	13	61.0%	7	40.0%	6	46.7%	12	41.7%
Term. Sac.	25	100.0%	23	100.0%	36	100.0%	32	100.0%	35	100.0%

Table 1B: Intercurrent Mortality Rate
Female Rats

	VEHICLE CONTROL		LOW		MEDIUM		HIGH1		PAIR-FED CONTROL	
Week	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT
0-52	7	11.7%	1	1.7%	3	5.0%	8	13.3%	1	1.7%
53-78	18	41.7%	15	26.7%	12	25.0%	10	30.0%	9	16.7%
79-92	8	55.0%	19	58.3%	7	36.7%	7	41.7%	16	43.3%
93-97	7	66.7%	1	60.0%	8	50.0%	6	51.7%	4	50.0%
Term. Sac.	20	100.0%	24	100.0%	30	100.0%	29	100.0%	30	100.0%

Table 2A: Intercurrent Mortality Comparison
Male Rats

Test	P-Value (across four groups)	P-Value (vehicle control vs low)	P-Value (vehicle control vs medium)	P-Value (vehicle control vs high)
Dose Response	0.4474	0.6574	0.2029	0.6044
Homogeneity	0.1585	0.5618	0.0887	0.4726

Test	P-Value (across four groups)	P-Value (pair-fed control vs low)	P-Value (pair-fed control vs medium)	P-Value (pair-fed control vs high)
Dose Response	0.9823	0.0951	0.9805	0.5455
Homogeneity	0.0758	0.0191	0.9187	0.3600

Table 2B: Intercurrent Mortality Comparison
Female Rats

Test	P-Value (across four groups)	P-Value (vehicle control vs low)	P-Value (vehicle control vs medium)	P-Value (vehicle control vs high)
Dose Response	0.2967	0.5134	0.1048	0.2382
Homogeneity	0.1592	0.4068	0.0320	0.1231

Test	P-Value (across four groups)	P-Value (pair-fed control vs low)	P-Value (pair-fed control vs medium)	P-Value (pair-fed control vs high)
Dose Response	0.9971	0.2670	0.9591	0.7119
Homogeneity	0.4994	0.1628	0.9548	0.6951

Table 3A (1): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (vehicle control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
ALL_SITES	HAEMANGIOSARCOMA+HAE	0	1	1	0	0.632	0.477	0.521	.
	LEUKEMIAS	2	1	3	0	0.877	0.862	0.530	1.000
	LYMPHOMAS	1	1	1	0	0.833	0.730	0.763	1.000
LIVER	HEP_ADENOMA+CARCINOM	1	2	0	0	0.941	0.474	1.000	1.000
SKIN_SUBCUTIS	BASAL_CELL_ADENOMA+C	3	1	1	0	0.970	0.927	0.947	1.000
	SQUAMOUS_CELL_PAPILL	5	2	1	0	0.997	0.921	0.988	1.000
THYMUS_GLAND	THYMOMA_BEGNIN+MALIG	1	1	0	0	0.937	0.729	1.000	1.000
THYROID	C_CELL_ADENOMA+CARCI	14	11	7	9	0.842	0.706	0.982	0.898
adrenal glands	ADENOMA, CORTICAL, B	1	0	0	0	1.000	1.000	1.000	1.000
	CARCINOMA, C-CELL, M	0	0	0	1	0.247	.	.	0.494
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.642	0.483	0.269	.
	LEUKEMIA, LARGE GRAN	0	0	1	0	0.517	.	0.516	.
	PHEOCHROMOCYTOMA, BE	6	6	5	2	0.949	0.520	0.751	0.964
	SCHWANNOMA, MALIGNAN	0	0	0	2	0.062	.	.	0.247
aorta	SARCOMA, HISTIOCYTIC	0	1	0	1	0.309	0.477	.	0.494
bone	LEUKEMIA, GRANULOCYT	0	0	1	0	0.520	.	0.521	.
bone marrow, fe	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LYMPHOMA, MALIGNANT	1	1	0	0	0.936	0.730	1.000	1.000
	SARCOMA, HISTIOCYTIC	1	0	0	0	1.000	1.000	1.000	1.000
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
bone marrow, st	HEMANGIOMA, BENIGN	0	0	1	0	0.520	.	0.521	.
	LEUKEMIA, GRANULOCYT	1	1	1	0	0.835	0.730	0.768	1.000
bone marrow, st	LYMPHOMA, MALIGNANT	1	1	1	0	0.833	0.730	0.763	1.000
	SARCOMA, HISTIOCYTIC	1	0	0	0	1.000	1.000	1.000	1.000
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
bone, femur	LEUKEMIA, GRANULOCYT	0	0	1	0	0.520	.	0.521	.
	SARCOMA, HISTIOCYTIC	1	0	0	0	1.000	1.000	1.000	1.000
bone, sternum	LEUKEMIA, GRANULOCYT	0	0	1	0	0.520	.	0.521	.
	PHEOCHROMOCYTOMA, MA	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	1	1	0	0	0.936	0.730	1.000	1.000
bone, vertebra	CHONDROSARCOMA, MALI	1	0	0	0	1.000	1.000	1.000	1.000

Table 3A (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (vehicle control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
brain	ASTROCYTOMA, MALIGNA	3	1	1	0	0.970	0.930	0.949	1.000
	CARCINOMA, PARS DIST	1	0	0	0	1.000	1.000	1.000	1.000
	CARCINOMA, SQUAMOUS	1	0	0	0	1.000	1.000	1.000	1.000
	GRANULAR CELL TUMOR,	0	1	0	0	0.747	0.477	.	.
	LEUKEMIA, GRANULOCYT	0	1	1	0	0.631	0.483	0.521	.
	LEUKEMIA, LARGE GRAN	0	0	1	0	0.517	.	0.516	.
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
cavity, abdomin	ADENOCARCINOMA (PRIM	0	1	0	0	0.749	0.483	.	.
	SARCOMA, HISTIOCYTIC	0	2	0	0	0.819	0.230	.	.
	SCHWANNOMA, MALIGNAN	0	0	0	2	0.062	.	.	0.247
cavity, oral	FIBROSARCOMA, MALIGN	0	0	0	1	0.247	.	.	0.494
cavity, thoraci	HIBERNOMA, MALIGNANT	1	1	5	2	0.334	0.723	0.117	0.492
	PHEOCHROMOCYTOMA, MA	1	0	0	0	1.000	1.000	1.000	1.000
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
coagulating gla	SARCOMA, HISTIOCYTIC	0	0	0	1	0.247	.	.	0.494
	SCHWANNOMA, MALIGNAN	0	0	0	1	0.247	.	.	0.494
ears	ADENOMA, BASAL CELL,	1	0	0	0	1.000	1.000	1.000	1.000
epididymides	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	1	0	0	0.749	0.483	.	.
	SCHWANNOMA, MALIGNAN	0	0	0	1	0.247	.	.	0.494
esophagus	SARCOMA, HISTIOCYTIC	0	1	0	1	0.309	0.477	.	0.494
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
eyes	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.247	.	.	0.494
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
eyes, optic ner	LEUKEMIA, GRANULOCYT	0	0	1	0	0.520	.	0.521	.
harderian gland	LEUKEMIA, GRANULOCYT	0	1	2	0	0.642	0.483	0.269	.
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
heart	ADENOCARCINOMA (PRIM	0	1	0	0	0.749	0.483	.	.
	LEUKEMIA, GRANULOCYT	1	0	2	0	0.697	1.000	0.524	1.000
	LEUKEMIA, LARGE GRAN	0	0	1	0	0.517	.	0.516	.
	LYMPHOMA, MALIGNANT	1	1	1	0	0.833	0.730	0.763	1.000

Table 3A (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (vehicle control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.247	.	.	0.494
	SCHWANNOMA, MALIGNAN	0	0	2	0	0.506	.	0.269	.
kidneys	ADENOMA, TUBULAR CEL	0	0	0	1	0.247	.	.	0.494
kidneys	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LIPOMA, BENIGN	0	0	1	0	0.517	.	0.516	.
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	RENAL MESENCHYMAL TU	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	1	0	2	0.108	0.483	.	0.247
	SCHWANNOMA, MALIGNAN	0	0	0	2	0.062	.	.	0.247
lacrimal glands	CARCINOMA, SQUAMOUS	0	1	0	0	0.747	0.477	.	.
	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
large intestine	LIPOMA, BENIGN	0	0	1	0	0.517	.	0.516	.
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.247	.	.	0.494
larynx	CARCINOMA, C-CELL, M	0	0	0	1	0.247	.	.	0.494
liver	ADENOMA, HEPATOCELLU	1	1	0	0	0.937	0.729	1.000	1.000
	CARCINOMA, HEPATOCEL	0	1	0	0	0.747	0.477	.	.
	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LEUKEMIA, LARGE GRAN	0	0	1	0	0.517	.	0.516	.
	LYMPHOMA, MALIGNANT	1	1	1	0	0.833	0.730	0.763	1.000
	SARCOMA, HISTIOCYTIC	1	1	0	2	0.236	0.723	1.000	0.492
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
lung	ADENOCARCINOMA (PRIM	0	1	0	0	0.749	0.483	.	.
	ADENOCARCINOMA, MALI	1	0	0	0	1.000	1.000	1.000	1.000
	CARCINOMA, C-CELL, M	0	0	0	1	0.247	.	.	0.494
	FIBROUS HISTIOCYTOMA	0	0	0	1	0.247	.	.	0.494
	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LEUKEMIA, LARGE GRAN	0	0	1	0	0.517	.	0.516	.
lung	LIPOSARCOMA (PRIMARY	0	0	1	0	0.517	.	0.516	.
	LYMPHANGIOSARCOMA, M	0	0	0	1	0.247	.	.	0.494
	LYMPHOMA, MALIGNANT	1	0	1	0	0.765	1.000	0.763	1.000
	OSTEOSARCOMA, MALIGN	1	0	0	0	1.000	1.000	1.000	1.000
	RENAL MESENCHYMAL TU	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	1	2	0	1	0.621	0.466	1.000	0.742
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.

Table 3A (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (vehicle control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
lymph node, axi	SARCOMA, HISTIOCYTIC	1	0	0	0	1.000	1.000	1.000	1.000
lymph node, cer	LEUKEMIA, GRANULOCYT	0	0	1	0	0.520	.	0.521	.
lymph node, ili	LEUKEMIA, GRANULOCYT	1	1	0	0	0.936	0.730	1.000	1.000
lymph node, ing	LEUKEMIA, GRANULOCYT	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	1	0	0	0	1.000	1.000	1.000	1.000
lymph node, man	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LYMPHOMA, MALIGNANT	1	0	1	0	0.765	1.000	0.763	1.000
	SARCOMA, HISTIOCYTIC	1	0	0	0	1.000	1.000	1.000	1.000
lymph node, med	LEUKEMIA, GRANULOCYT	0	0	1	0	0.520	.	0.521	.
lymph node, mes	HEMANGIOSARCOMA, MAL	0	1	0	0	0.747	0.477	.	.
	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LYMPHOMA, MALIGNANT	1	1	1	0	0.833	0.730	0.763	1.000
	SARCOMA, HISTIOCYTIC	1	1	0	1	0.523	0.730	1.000	0.742
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
lymph node, ren	LEUKEMIA, GRANULOCYT	1	0	0	0	1.000	1.000	1.000	1.000
mammary gland	ADENOCARCINOMA, MALI	2	1	1	1	0.677	0.857	0.887	0.871
	FIBROADENOMA, BENIGN	1	0	1	0	0.765	1.000	0.763	1.000
multicentric ne	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LEUKEMIA, LARGE GRAN	1	0	1	0	0.768	1.000	0.769	1.000
	LYMPHOMA, MALIGNANT	1	1	1	0	0.833	0.730	0.763	1.000
	SARCOMA, HISTIOCYTIC	1	2	0	2	0.332	0.466	1.000	0.492
nerve, sciatic	LEUKEMIA, GRANULOCYT	0	0	1	0	0.520	.	0.521	.
pancreas	ADENOCARCINOMA (PRIM	0	1	0	0	0.749	0.483	.	.
	ADENOMA, ISLET CELL,	7	6	9	3	0.906	0.662	0.447	0.948
	CARCINOMA, ACINAR CE	0	0	1	0	0.517	.	0.516	.
	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LYMPHOMA, MALIGNANT	1	0	1	0	0.765	1.000	0.763	1.000
	SARCOMA, HISTIOCYTIC	1	0	0	1	0.432	1.000	1.000	0.742
	SCHWANNOMA, MALIGNAN	0	0	0	1	0.247	.	.	0.494
parathyroid gla	ADENOMA, BENIGN	2	1	0	0	0.985	0.861	1.000	1.000
	LEUKEMIA, GRANULOCYT	0	0	1	0	0.520	.	0.521	.

Table 3A (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (vehicle control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
pituitary gland	ADENOMA, PARS DISTAL	37	26	23	16	1.000	0.943	0.996	1.000
	ADENOMA, PARS INTERM	0	0	1	0	0.517	.	0.516	.
	CARCINOMA, PARS DIST	1	0	0	0	1.000	1.000	1.000	1.000
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.642	0.483	0.269	.
	LYMPHOMA, MALIGNANT	0	0	1	0	0.517	.	0.516	.
preputial gland	CARCINOMA, SQUAMOUS	1	0	0	0	1.000	1.000	1.000	1.000
prostate gland	LEUKEMIA, GRANULOCYT	0	1	0	0	0.749	0.483	.	.
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.247	.	.	0.494
prostate gland	SCHWANNOMA, MALIGNAN	0	0	1	1	0.194	.	0.516	0.494
salivary gland,	LEUKEMIA, GRANULOCYT	0	0	1	0	0.520	.	0.521	.
			1	2	0	0.642	0.483	0.269	.
	LEUKEMIA, LARGE GRAN	0	0	1	0	0.517	.	0.516	.
seminal vesicle	LEUKEMIA, GRANULOCYT	0	0	2	0	0.506	.	0.269	.
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.247	.	.	0.494
	SCHWANNOMA, MALIGNAN	0	0	0	1	0.247	.	.	0.494
skeletal muscle	LEUKEMIA, GRANULOCYT	0	0	2	0	0.506	.	0.269	.
			1	0	0	0.749	0.483	.	.
	LYMPHOMA, MALIGNANT	0	0	1	0	0.517	.	0.516	.
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.247	.	.	0.494
skin / skin, su	ADENOMA, BASAL CELL,	3	0	0	0	1.000	1.000	1.000	1.000
	CARCINOMA, BASAL CEL	0	1	1	0	0.631	0.477	0.516	.
	CARCINOMA, SQUAMOUS	2	0	0	0	1.000	1.000	1.000	1.000
	FIBROMA, BENIGN	1	1	1	0	0.838	0.729	0.769	1.000
	FIBROUS HISTIOCYTOMA	0	1	0	1	0.309	0.477	.	0.494
	KERATOACANTHOMA, BEN	2	2	1	0	0.948	0.647	0.887	1.000
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.642	0.483	0.269	.
	LYMPHANGIOSARCOMA, M	0	0	0	1	0.247	.	.	0.494
	OSTEOSARCOMA, MALIGN	1	1	0	0	0.935	0.723	1.000	1.000
	PAPILLOMA, SQUAMOUS	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	1	1	0	1	0.525	0.723	1.000	0.742
	SCHWANNOMA, MALIGNAN	0	1	1	0	0.631	0.477	0.516	.
small intestine	ADENOCARCINOMA, MALI	0	1	0	0	0.747	0.477	.	.
	LEUKEMIA, GRANULOCYT	0	0	1	0	0.520	.	0.521	.
	SARCOMA, HISTIOCYTIC	0	1	0	0	0.749	0.483	.	.
spinal cord, ce	LEUKEMIA, LARGE GRAN	0	0	1	0	0.517	.	0.516	.
spinal cord, lu	LEUKEMIA, LARGE GRAN	0	0	1	0	0.517	.	0.516	.

Table 3A (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (vehicle control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
spinal cord, th	LEUKEMIA, LARGE GRAN	0	0	1	0	0.517	.	0.516	.
spleen	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LEUKEMIA, LARGE GRAN	1	0	1	0	0.768	1.000	0.769	1.000
	LYMPHOMA, MALIGNANT	1	1	1	0	0.833	0.730	0.763	1.000
	SARCOMA, HISTIOCYTIC	1	0	0	0	1.000	1.000	1.000	1.000
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
stomach, glandu	ADENOCARCINOMA (PRIM	0	1	0	0	0.749	0.483	.	.
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.642	0.483	0.269	.
	POLYP, BENIGN	0	0	1	0	0.517	.	0.516	.
	SARCOMA, HISTIOCYTIC	1	0	0	1	0.432	1.000	1.000	0.742
	SCHWANNOMA, MALIGNAN	0	0	1	1	0.194	.	0.516	0.494
stomach, nongla	SARCOMA, HISTIOCYTIC	0	0	0	1	0.247	.	.	0.494
testes	ADENOMA, INTERSTITIA	5	2	2	8	0.039	0.921	0.948	0.234
	LEUKEMIA, GRANULOCYT	0	0	1	0	0.520	.	0.521	.
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	1	0	1	0.308	0.483	.	0.494
thymus gland	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LYMPHOMA, MALIGNANT	1	1	1	0	0.833	0.730	0.763	1.000
	SARCOMA, HISTIOCYTIC	0	1	0	1	0.308	0.483	.	0.494
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
	THYMOMA, BENIGN	1	0	0	0	1.000	1.000	1.000	1.000
	THYMOMA, MALIGNANT	0	1	0	0	0.747	0.477	.	.
thyroid gland	ADENOMA, C-CELL, BEN	14	10	7	8	0.887	0.790	0.982	0.939
	ADENOMA, FOLLICULAR	0	0	1	0	0.517	.	0.516	.
	CARCINOMA, C-CELL, M	0	1	0	1	0.309	0.477	.	0.494
	LEUKEMIA, GRANULOCYT	1	1	2	0	0.794	0.730	0.524	1.000
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.247	.	.	0.494
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
tongue	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
trachea	ADENOCARCINOMA, MALI	1	0	0	0	1.000	1.000	1.000	1.000
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.517	.	0.516	.
urinary bladder	LEIOMYOMA, BENIGN	0	0	0	1	0.247	.	.	0.494
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.642	0.483	0.269	.
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	1	0	1	0.308	0.483	.	0.494
	SCHWANNOMA, MALIGNAN	0	0	0	1	0.247	.	.	0.494

Table 3A (2): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (pair-fed control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
ALL_SITES	HAEMANGIOSARCOMA+HAE	1	1	1	0	0.822	0.701	0.748	1.000
	LEUKEMIAS	1	1	3	0	0.786	0.707	0.301	1.000
	LYMPHOMAS	2	1	1	0	0.919	0.844	0.871	1.000
LIVER	HEP_ADENOMA+CARCINOM	0	2	0	0	0.803	0.206	.	.
SKIN_SUBCUTIS	BASAL_CELL_ADENOMA+C	0	1	1	0	0.610	0.451	0.490	.
	SQUAMOUS_CELL_PAPILL	1	2	1	0	0.867	0.425	0.742	1.000
THYMUS_GLAND	THYMOMA_BEGNIN+MALIG	0	1	0	0	0.727	0.451	.	.
THYROID	C_CELL_ADENOMA+CARCI	13	11	7	9	0.757	0.536	0.952	0.795
adrenal glands	CARCINOMA, C-CELL, M	0	0	0	1	0.240	.	.	0.468
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LEUKEMIA, LARGE GRAN	1	0	1	0	0.754	1.000	0.742	1.000
	PHEOCHROMOCYTOMA, BE	8	6	5	2	0.975	0.679	0.867	0.987
	SCHWANNOMA, MALIGNAN	0	0	0	2	0.059	.	.	0.222
aorta	SARCOMA, HISTIOCYTIC	0	1	0	1	0.292	0.451	.	0.468
bone	LEUKEMIA, GRANULOCYT	0	0	1	0	0.505	.	0.495	.
bone marrow, fe	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LYMPHOMA, MALIGNANT	0	1	0	0	0.728	0.457	.	.
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
bone marrow, st	HEMANGIOMA, BENIGN	0	0	1	0	0.505	.	0.495	.
	LEUKEMIA, GRANULOCYT	0	1	1	0	0.610	0.457	0.495	.
	LYMPHOMA, MALIGNANT	0	1	1	0	0.609	0.457	0.490	.
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
bone, femur	LEUKEMIA, GRANULOCYT	0	0	1	0	0.505	.	0.495	.
bone, sternum	LEUKEMIA, GRANULOCYT	0	0	1	0	0.505	.	0.495	.
	SARCOMA, HISTIOCYTIC	0	1	0	0	0.728	0.457	.	.
brain	ASTROCYTOMA, MALIGNA	2	1	1	0	0.921	0.844	0.875	1.000
	GRANULAR CELL TUMOR,	0	1	0	0	0.727	0.451	.	.
	LEUKEMIA, GRANULOCYT	0	1	1	0	0.610	0.457	0.495	.
	LEUKEMIA, LARGE GRAN	1	0	1	0	0.754	1.000	0.742	1.000
cavity, abdomin	ADENOCARCINOMA (PRIM	0	1	0	0	0.728	0.457	.	.
	SARCOMA, HISTIOCYTIC	0	2	0	0	0.803	0.206	.	.

Table 3A (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (pair-fed control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
	SCHWANNOMA, MALIGNAN	0	0	0	2	0.059	.	.	0.222
cavity, oral	FIBROSARCOMA, MALIGN	0	0	0	1	0.240	.	.	0.468
cavity, thoraci	HIBERNOMA_BENIGN+MAL	1	1	5	2	0.313	0.701	0.098	0.460
	HIBERNOMA, BENIGN	0	0	0	2	0.057	.	.	0.226
	HIBERNOMA, MALIGNANT	0	1	1	1	0.284	0.472	0.500	0.478
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
coagulating gla	SARCOMA, HISTIOCYTIC	0	0	0	1	0.240	.	.	0.468
	SCHWANNOMA, MALIGNAN	0	0	0	1	0.240	.	.	0.468
epididymides	SARCOMA, HISTIOCYTIC	0	1	0	0	0.728	0.457	.	.
	SCHWANNOMA, MALIGNAN	0	0	0	1	0.240	.	.	0.468
esophagus	SARCOMA, HISTIOCYTIC	0	1	0	1	0.292	0.451	.	0.468
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
eyes	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.245	.	.	0.474
eyes	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
eyes, optic ner	LEUKEMIA, GRANULOCYT	0	0	1	0	0.505	.	0.495	.
harderian gland	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
heart	ADENOCARCINOMA (PRIM	0	1	0	0	0.728	0.457	.	.
	LEUKEMIA, GRANULOCYT	0	0	2	0	0.492	.	0.242	.
	LEUKEMIA, LARGE GRAN	1	0	1	0	0.754	1.000	0.742	1.000
	LYMPHOMA, MALIGNANT	1	1	1	0	0.820	0.707	0.742	1.000
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.240	.	.	0.468
	SCHWANNOMA, MALIGNAN	0	0	2	0	0.492	.	0.242	.
kidneys	ADENOMA, TUBULAR CEL	0	0	0	1	0.240	.	.	0.468
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LEUKEMIA, LARGE GRAN	1	0	0	0	1.000	1.000	1.000	1.000
	LIPOMA, BENIGN	0	0	1	0	0.503	.	0.490	.
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	1	0	2	0.100	0.457	.	0.222
	SCHWANNOMA, MALIGNAN	0	0	0	2	0.059	.	.	0.222
lacrimal glands	CARCINOMA, SQUAMOUS	0	1	0	0	0.727	0.451	.	.
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000

Table 3A (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (pair-fed control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
large intestine	LIPOMA, BENIGN	0	0	1	0	0.503	.	0.490	.
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.240	.	.	0.468
larynx	CARCINOMA, C-CELL, M	0	0	0	1	0.240	.	.	0.468
liver	ADENOMA, HEPATOCELLU	0	1	0	0	0.727	0.451	.	.
	CARCINOMA, HEPATOCEL	0	1	0	0	0.727	0.451	.	.
	CHOLANGIOMA, BENIGN	1	0	0	0	1.000	1.000	1.000	1.000
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LEUKEMIA, LARGE GRAN	1	0	1	0	0.754	1.000	0.742	1.000
	LYMPHOMA, MALIGNANT	1	1	1	0	0.820	0.707	0.742	1.000
	SARCOMA, HISTIOCYTIC	0	1	0	2	0.100	0.451	.	0.222
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
lung	ADENOCARCINOMA (PRIM	0	1	0	0	0.728	0.457	.	.
	CARCINOMA, C-CELL, M	0	0	0	1	0.240	.	.	0.468
	FIBROUS HISTIOCYTOMA	0	0	0	1	0.240	.	.	0.468
	HEMANGIOSARCOMA, MAL	1	0	0	0	1.000	1.000	1.000	1.000
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LEUKEMIA, LARGE GRAN	1	0	1	0	0.754	1.000	0.742	1.000
	LIPOSARCOMA (PRIMARY	0	0	1	0	0.505	.	0.495	.
	LYMPHANGIOSARCOMA, M	0	0	0	1	0.240	.	.	0.468
	LYMPHOMA, MALIGNANT	1	0	1	0	0.754	1.000	0.742	1.000
	SARCOMA, HISTIOCYTIC	0	2	0	1	0.419	0.206	.	0.468
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
lymph node, cer	LEUKEMIA, GRANULOCYT	0	0	1	0	0.505	.	0.495	.
lymph node, hep	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
lymph node, ili	LEUKEMIA, GRANULOCYT	0	1	0	0	0.728	0.457	.	.
lymph node, man	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LYMPHOMA, MALIGNANT	0	0	1	0	0.503	.	0.490	.
lymph node, med	LEUKEMIA, GRANULOCYT	0	0	1	0	0.505	.	0.495	.
lymph node, med	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
lymph node, mes	HEMANGIOSARCOMA, MAL	0	1	0	0	0.727	.	.	.
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LYMPHOMA, MALIGNANT	1	1	1	0	0.820	0.707	0.742	1.000
	SARCOMA, HISTIOCYTIC	0	1	0	1	0.296	0.457	.	0.474
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.

Table 3A (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (pair-fed control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
mammary gland	ADENOCARCINOMA, MALI	0	1	1	1	0.280	0.451	0.495	0.468
	FIBROADENOMA, BENIGN	0	0	1	0	0.505	.	0.495	.
mediastinum	SARCOMA, HISTIOCYTIC	1	0	0	0	1.000	1.000	1.000	1.000
	SCHWANNOMA, MALIGNAN	1	0	0	0	1.000	1.000	1.000	1.000
multicentric ne	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LEUKEMIA, LARGE GRAN	1	0	1	0	0.754	1.000	0.742	1.000
	LYMPHOMA, MALIGNANT	2	1	1	0	0.919	0.844	0.871	1.000
	SARCOMA, HISTIOCYTIC	3	2	0	2	0.599	0.760	1.000	0.785
nerve, sciatic	LEUKEMIA, GRANULOCYT	0	0	1	0	0.505	.	0.495	.
	LEUKEMIA, LARGE GRAN	1	0	0	0	1.000	1.000	1.000	1.000
pancreas	ADENOCARCINOMA (PRIM	0	1	0	0	0.728	0.457	.	.
	ADENOMA, ISLET CELL,	4	6	9	3	0.729	0.253	0.102	0.717
	CARCINOMA, ACINAR CE	0	0	1	0	0.505	.	0.495	.
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LYMPHOMA, MALIGNANT	1	0	1	0	0.754	1.000	0.742	1.000
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.240	.	.	0.468
	SCHWANNOMA, MALIGNAN	0	0	0	1	0.240	.	.	0.468
parathyroid gla	ADENOMA, BENIGN	2	1	0	0	0.981	0.839	1.000	1.000
parathyroid gla	LEUKEMIA, GRANULOCYT	0	0	1	0	0.505	.	0.495	.
pituitary gland	ADENOMA, PARS DISTAL	20	26	23	16	0.890	0.034	0.211	0.701
	ADENOMA, PARS INTERM	0	0	1	0	0.505	.	0.495	.
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LYMPHOMA, MALIGNANT	0	0	1	0	0.503	.	0.490	.
prostate gland	LEUKEMIA, GRANULOCYT	0	1	0	0	0.728	0.457	.	.
	LEUKEMIA, LARGE GRAN	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.240	.	.	0.468
	SCHWANNOMA, MALIGNAN	0	0	1	1	0.184	.	0.490	0.468
salivary gland,	LEUKEMIA, GRANULOCYT	0	0	1	0	0.505	.	0.495	.
			1	2	0	0.625	0.457	0.242	.
	LEUKEMIA, LARGE GRAN	0	0	1	0	0.503	.	0.490	.
	SCHWANNOMA, MALIGNAN	1	0	0	0	1.000	1.000	1.000	1.000
seminal vesicle	LEUKEMIA, GRANULOCYT	0	0	2	0	0.492	.	0.242	.
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.240	.	.	0.468
	SCHWANNOMA, MALIGNAN	0	0	0	1	0.240	.	.	0.468

Table 3A (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (pair-fed control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
skeletal muscle	LEUKEMIA, GRANULOCYT	0	0	2	0	0.492	.	0.242	.
			1	0	0	0.728	0.457	.	.
	LYMPHOMA, MALIGNANT	0	0	1	0	0.503	.	0.490	.
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.240	.	.	0.468
skin	CARCINOMA, BASAL CEL	0	1	1	0	0.610	0.451	0.490	.
	KERATOACANTHOMA, BEN	1	2	1	0	0.867	0.425	0.742	1.000
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.240	.	.	0.468
	TRICHOEPITHELIOMA, B	1	0	0	0	1.000	1.000	1.000	1.000
skin, subcutis	FIBROMA, BENIGN	0	1	1	0	0.610	0.451	0.490	.
	FIBROSARCOMA, MALIGN	1	0	0	0	1.000	1.000	1.000	1.000
	FIBROUS HISTIOCYTOMA	0	1	0	1	0.292	0.451	.	0.468
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LYMPHANGIOSARCOMA, M	0	0	0	1	0.240	.	.	0.468
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	OSTEOSARCOMA, MALIGN	0	1	0	0	0.727	0.451	.	.
	SARCOMA, HISTIOCYTIC	2	1	0	1	0.672	0.839	1.000	0.854
	SCHWANNOMA, MALIGNAN	2	1	1	0	0.920	0.839	0.871	1.000
small intestine	ADENOCARCINOMA, MALI	0	1	0	0	0.727	0.451	.	.
	LEUKEMIA, GRANULOCYT	0	0	1	0	0.505	.	0.495	.
	SARCOMA, HISTIOCYTIC	0	1	0	0	0.728	0.457	.	.
spinal cord, ce	LEUKEMIA, LARGE GRAN	0	0	1	0	0.503	.	0.490	.
spinal cord, lu	LEUKEMIA, LARGE GRAN	0	0	1	0	0.503	.	0.490	.
spinal cord, th	LEUKEMIA, LARGE GRAN	1	0	1	0	0.754	1.000	0.742	1.000
spleen	HEMANGIOSARCOMA, MAL	1	0	0	0	1.000	1.000	1.000	1.000
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LEUKEMIA, LARGE GRAN	1	0	1	0	0.754	1.000	0.742	1.000
	LYMPHOMA, MALIGNANT	1	1	1	0	0.820	0.707	0.742	1.000
	SARCOMA, HISTIOCYTIC	1	0	0	0	1.000	1.000	1.000	1.000
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
stomach, glandu	ADENOCARCINOMA (PRIM	0	1	0	0	0.728	0.457	.	.
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	POLYP, BENIGN	0	0	1	0	0.503	.	0.490	.
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.240	.	.	0.468
stomach, glandu	SCHWANNOMA, MALIGNAN	0	0	1	1	0.184	.	0.490	0.468

Table 3A (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (pair-fed control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=59	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
stomach, nongla	SARCOMA, HISTIOCYTIC	0	0	0	1	0.240	.	.	0.468
testes	ADENOMA, INTERSTITIA	4	2	2	8	0.019	0.846	0.888	0.122
	LEUKEMIA, GRANULOCYT	0	0	1	0	0.505	.	0.495	.
	LEUKEMIA, LARGE GRAN	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	1	0	1	0.291	0.457	.	0.468
thymus gland	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LYMPHOMA, MALIGNANT	2	1	1	0	0.919	0.844	0.871	1.000
	SARCOMA, HISTIOCYTIC	0	1	0	1	0.291	0.457	.	0.468
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
	THYMOMA, MALIGNANT	0	1	0	0	0.727	0.451	.	.
thyroid gland	ADENOMA, C-CELL, BEN	13	10	7	8	0.817	0.640	0.952	0.865
	ADENOMA, FOLLICULAR	0	0	1	0	0.503	.	0.490	.
	CARCINOMA, C-CELL, M	0	1	0	1	0.292	0.451	.	0.468
	CARCINOMA, FOLLICULA	1	0	0	0	1.000	1.000	1.000	1.000
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	SARCOMA, HISTIOCYTIC	0	0	0	1	0.240	.	.	0.468
	SCHWANNOMA, MALIGNAN	0	0	1	0	0.503	.	0.490	.
tongue	SCHWANNOMA, MALIGNAN	1	0	1	0	0.754	1.000	0.742	1.000
trachea	SCHWANNOMA, MALIGNAN	1	0	1	0	0.754	1.000	0.742	1.000
urinary bladder	LEIOMYOMA, BENIGN	0	0	0	1	0.240	.	.	0.468
	LEUKEMIA, GRANULOCYT	0	1	2	0	0.625	0.457	0.242	.
	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	1	0	1	0.291	0.457	.	0.468
urinary bladder	SCHWANNOMA, MALIGNAN	0	0	0	1	0.240	.	.	0.468

Table 3B (1): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (vehicle control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=60	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
ALL_SITES	LYMPHOMAS	1	0	0	0	1.000	1.000	1.000	1.000
SKIN_subcutis	CARCINOMA+KERATOACAN	0	0	2	1	0.218	.	0.309	0.531
THYROID	FOLLICULAR_CELL_ADEN	1	0	0	1	0.441	1.000	1.000	0.778
adrenal glands	ADENOMA, CORTICAL, B	0	2	1	0	0.760	0.266	0.548	.
	CARCINOMA, CORTICAL,	0	0	0	1	0.256	.	.	0.531
	PHEOCHROMOCYTOMA, BE	0	2	1	0	0.761	0.266	0.553	.
brain	ASTROCYTOMA, MALIGNA	1	0	1	4	0.028	1.000	0.803	0.237
	CARCINOMA, PARS DIST	0	2	1	0	0.760	0.266	0.548	.
	GRANULAR CELL TUMOR,	0	1	0	0	0.773	0.519	.	.
	PINEALOMA, MALIGNANT	0	0	1	0	0.530	.	0.553	.
cavity, thoraci	HIBERNOMA, MALIGNANT	1	1	1	3	0.119	0.778	0.803	0.347
cervix	FIBROSARCOMA, MALIGN	0	1	0	0	0.774	0.525	.	.
clitoral glands	ADENOCARCINOMA, MALI	0	0	1	0	0.530	.	0.553	.
heart	ADENOCARCINOMA, MALI	0	0	1	0	0.527	.	0.548	.
kidneys	ADENOMA, TUBULAR CEL	0	0	0	1	0.256	.	.	0.531
	LIPOMA, BENIGN	0	1	0	0	0.774	0.525	.	.
liver	ADENOCARCINOMA, MALI	1	0	0	0	1.000	1.000	1.000	1.000
	CARCINOMA, CORTICAL,	0	0	0	1	0.256	.	.	0.531
	CHOLANGIOMA, BENIGN	1	0	0	0	1.000	1.000	1.000	1.000
lung	ADENOCARCINOMA, MALI	0	2	1	1	0.438	0.279	0.548	0.525
lung	CARCINOMA, CORTICAL,	0	0	0	1	0.256	.	.	0.531
	HIBERNOMA, MALIGNANT	0	0	1	0	0.530	.	0.553	.
lymph node, axi	ADENOCARCINOMA, MALI	0	1	0	0	0.774	0.525	.	.
lymph node, med	ADENOCARCINOMA, MALI	0	1	0	0	0.774	0.525	.	.
mammary gland	ADENOCARCINOMA, MALI	16	23	18	20	0.385	0.206	0.600	0.318
	ADENOMA, BENIGN	1	0	0	0	1.000	1.000	1.000	1.000
	FIBROADENOMA, BENIGN	21	29	20	9	1.000	0.130	0.832	0.998

Table 3B (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (vehicle control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=60	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
multicentric ne	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	0	1	0	0.530	.	0.553	.
ovaries	GRANULOSA CELL TUMOR	0	0	1	0	0.527	.	0.548	.
	SERTOLI CELL TUMOR,	1	0	0	0	1.000	1.000	1.000	1.000
pancreas	ADENOMA, ISLET CELL,	1	1	2	1	0.523	0.766	0.570	0.771
parathyroid gla	ADENOMA, BENIGN	1	0	0	0	1.000	1.000	1.000	1.000
pituitary gland	ADENOMA, PARS DISTAL	46	43	36	32	0.981	0.955	0.996	0.997
	ADENOMA, PARS INTERM	0	0	0	1	0.252	.	.	0.525
	CARCINOMA, PARS DIST	0	2	1	0	0.760	0.266	0.548	.
skin / skin, su	ADENOMA, BASAL CELL,	0	1	0	0	0.773	0.519	.	.
	CARCINOMA, SQUAMOUS	0	0	1	1	0.206	.	0.553	0.531
	FIBROMA, BENIGN	1	0	0	1	0.439	1.000	1.000	0.771
	FIBROUS HISTIOCYTOMA	1	0	0	0	1.000	1.000	1.000	1.000
	KERATOACANTHOMA, BEN	0	0	1	0	0.530	.	0.553	.
	SARCOMA, HISTIOCYTIC	0	0	1	0	0.530	.	0.553	.
spinal cord, ce	ASTROCYTOMA, MALIGNA	0	0	0	1	0.252	.	.	0.525
stomach, nongla	PAPILLOMA, SQUAMOUS	0	0	0	1	0.252	.	.	0.525
thymus gland	LYMPHOMA, MALIGNANT	1	0	0	0	1.000	1.000	1.000	1.000
thyroid gland	ADENOMA, C-CELL, BEN	5	5	7	5	0.539	0.678	0.517	0.678
	ADENOMA, FOLLICULAR	0	0	0	1	0.252	.	.	0.525
	CARCINOMA, FOLLICULA	1	0	0	0	1.000	1.000	1.000	1.000
uterus with cer	GRANULAR CELL TUMOR,	0	0	3	5	0.005	.	0.164	0.038
	HEMANGIOSARCOMA, MAL	1	0	0	0	1.000	1.000	1.000	1.000
	POLYP, STROMAL, BENI	0	1	4	3	0.100	0.519	0.088	0.140
vagina	GRANULAR CELL TUMOR,	0	1	0	2	0.112	0.525	.	0.273
	LEIOMYOMA, BENIGN	1	0	0	1	0.445	1.000	1.000	0.777

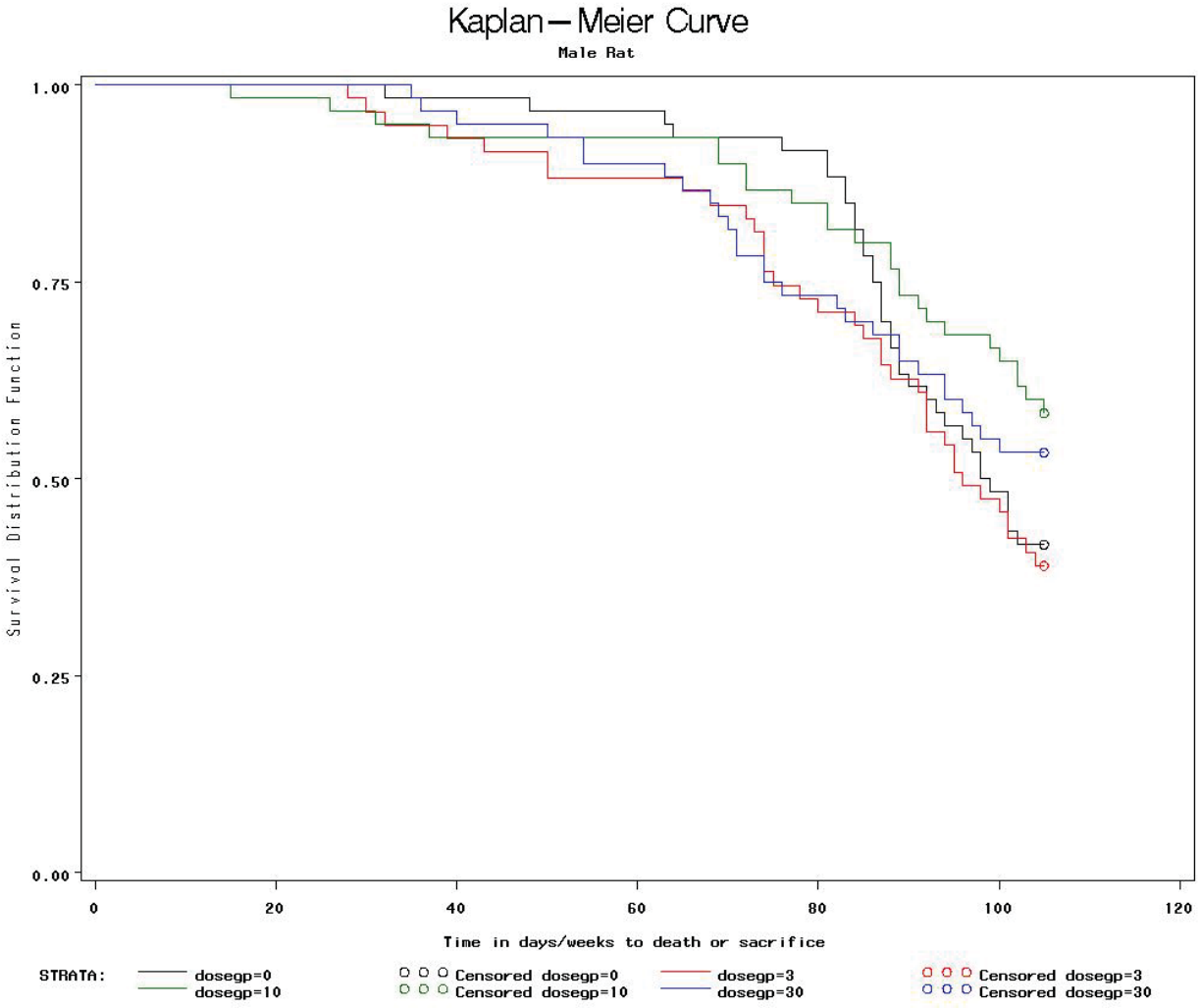
Table 3B (2): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (pair-fed control, low, medium and high dose groups)

Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=60	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
SKIN_subcutis	CARCINOMA+KERATOACAN	0	0	2	1	0.195	.	0.253	0.478
THYROID	FOLLICULAR_CELL_ADEN	0	0	0	1	0.242	.	.	0.478
adrenal glands	ADENOMA, CORTICAL, B	3	2	1	0	0.975	0.767	0.939	1.000
	CARCINOMA, CORTICAL,	0	0	0	1	0.242	.	.	0.478
	PHEOCHROMOCYTOMA, BE	0	2	1	0	0.726	0.220	0.500	.
brain	ASTROCYTOMA, MALIGNA	3	0	1	4	0.103	1.000	0.939	0.464
	CARCINOMA, PARS DIST	0	2	1	0	0.728	0.214	0.500	.
	GRANULAR CELL TUMOR,	0	1	0	0	0.736	0.466	.	.
	PINEALOMA, MALIGNANT	0	0	1	0	0.506	.	0.500	.
cavity, cranial	SARCOMA, HISTIOCYTIC	1	0	0	0	1.000	1.000	1.000	1.000
cavity, thoraci	HIBERNOMA, MALIGNANT	0	1	1	3	0.038	0.472	0.500	0.105
cervix	FIBROSARCOMA, MALIGN	0	1	0	0	0.737	0.472	.	.
clitoral glands	ADENOCARCINOMA, MALI	0	0	1	0	0.506	.	0.500	.
ears	SCHWANNOMA, MALIGNAN	2	0	0	0	1.000	1.000	1.000	1.000
heart	ADENOCARCINOMA, MALI	0	0	1	0	0.506	.	0.500	.
kidneys	ADENOMA, TUBULAR CEL	0	0	0	1	0.242	.	.	0.478
	LIPOMA, BENIGN	0	1	0	0	0.737	0.472	.	.
liver	CARCINOMA, CORTICAL,	0	0	0	1	0.242	.	.	0.478
lung	ADENOCARCINOMA, MALI	0	2	1	1	0.400	0.226	0.500	0.478
lung	CARCINOMA, CORTICAL,	0	0	0	1	0.242	.	.	0.478
	HIBERNOMA, MALIGNANT	0	0	1	0	0.506	.	0.500	.
lymph node, axi	ADENOCARCINOMA, MALI	0	1	0	0	0.737	0.472	.	.
lymph node, med	ADENOCARCINOMA, MALI	0	1	0	0	0.737	0.472	.	.
	CARCINOMA, C-CELL, M	1	0	0	0	1.000	1.000	1.000	1.000
mammary gland	ADENOCARCINOMA, MALI	20	23	18	20	0.383	0.207	0.623	0.326
	FIBROADENOMA, BENIGN	24	29	20	9	1.000	0.106	0.820	0.999
multicentric ne	SARCOMA, HISTIOCYTIC	1	0	1	0	0.754	1.000	0.747	1.000

Table 3B (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (pair-fed control, low, medium and high dose groups)

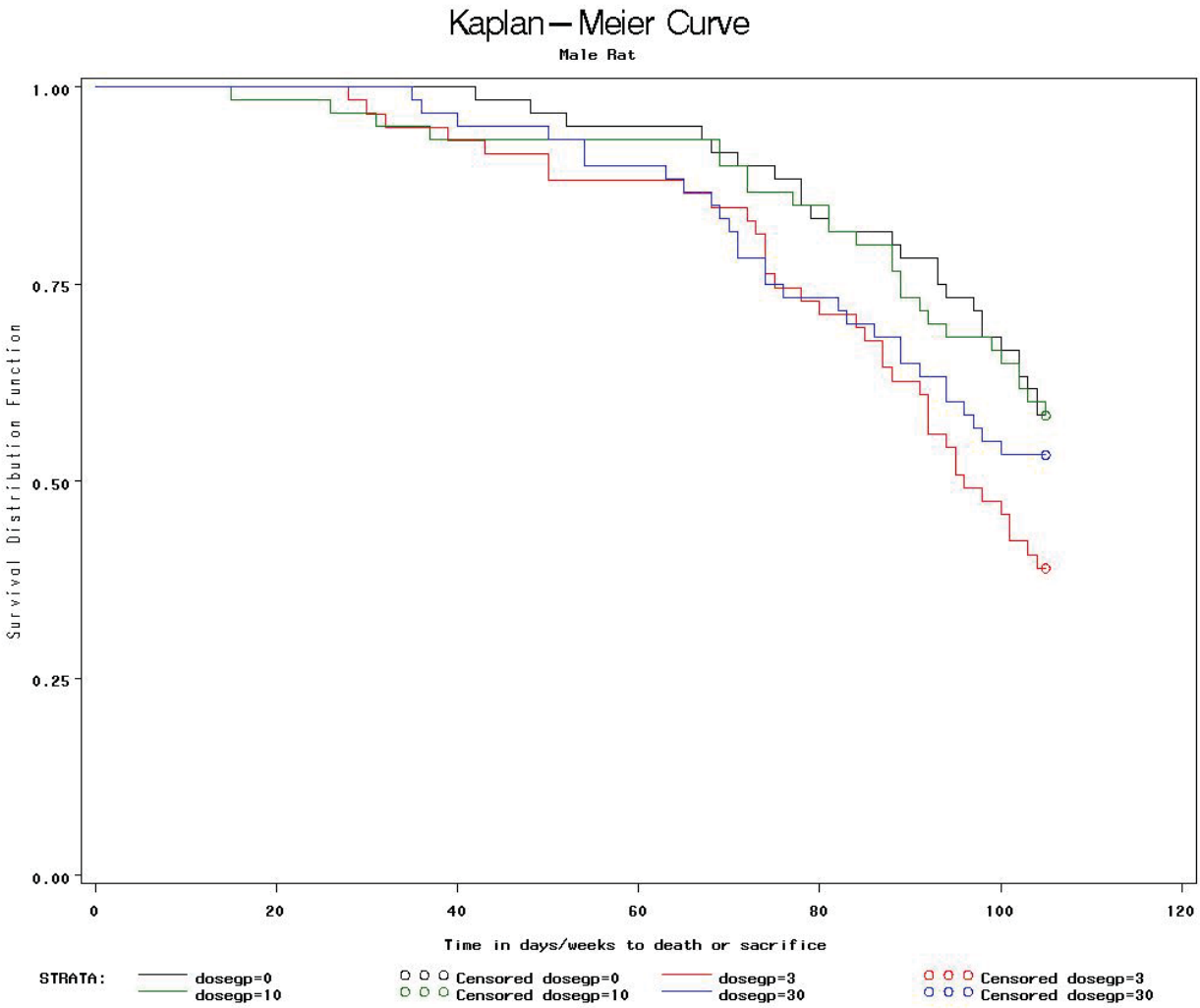
Organ Name	Tumor Name	0 mg Cont N=60	3 mg Low N=60	10 mg Med N=60	30 mg High N=60	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
ovaries	GRANULOSA CELL TUMOR	0	0	1	0	0.506	.	0.500	.
pancreas	ADENOMA, ISLET CELL,	2	1	2	1	0.640	0.852	0.692	0.862
	CARCINOMA, ISLET CEL	1	0	0	0	1.000	1.000	1.000	1.000
pituitary gland	ADENOMA, PARS DISTAL	47	43	36	32	0.939	0.729	0.952	0.960
	ADENOMA, PARS INTERM	0	0	0	1	0.242	.	.	0.478
	CARCINOMA, PARS DIST	1	2	1	0	0.869	0.441	0.747	1.000
skin	ADENOMA, BASAL CELL,	0	1	0	0	0.736	0.466	.	.
	CARCINOMA, SQUAMOUS	0	0	1	1	0.186	.	0.500	0.478
	KERATOACANTHOMA, BEN	0	0	1	0	0.506	.	0.500	.
skin, subcutis	FIBROMA, BENIGN	1	0	0	1	0.426	1.000	1.000	0.730
	FIBROSARCOMA, MALIGN	1	0	0	0	1.000	1.000	1.000	1.000
	FIBROUS HISTIOCYTOMA	1	0	0	0	1.000	1.000	1.000	1.000
	SARCOMA, HISTIOCYTIC	0	0	1	0	0.506	.	0.500	.
small intestine	LEIOMYOSARCOMA, MALI	1	0	0	0	1.000	1.000	1.000	1.000
spinal cord, ce	ASTROCYTOMA, MALIGNA	0	0	0	1	0.242	.	.	0.478
stomach, nongla	PAPILLOMA, SQUAMOUS	0	0	0	1	0.242	.	.	0.478
thyroid gland	ADENOMA, C-CELL, BEN	9	5	7	5	0.751	0.883	0.781	0.892
	ADENOMA, FOLLICULAR	0	0	0	1	0.242	.	.	0.478
	CARCINOMA, C-CELL, M	1	0	0	0	1.000	1.000	1.000	1.000
uterus with cer	GRANULAR CELL TUMOR,	3	0	3	5	0.050	1.000	0.651	0.297
	POLYP, STROMAL, BENI	4	1	4	3	0.421	0.959	0.631	0.735
vagina	GRANULAR CELL TUMOR,	0	1	0	2	0.099	0.472	.	0.226
	LEIOMYOMA, BENIGN	0	0	0	1	0.242	.	.	0.478

Figure 1A (1): Kaplan-Meier Survival Functions for Male Rats
(vehicle control, low, medium and high dose groups)



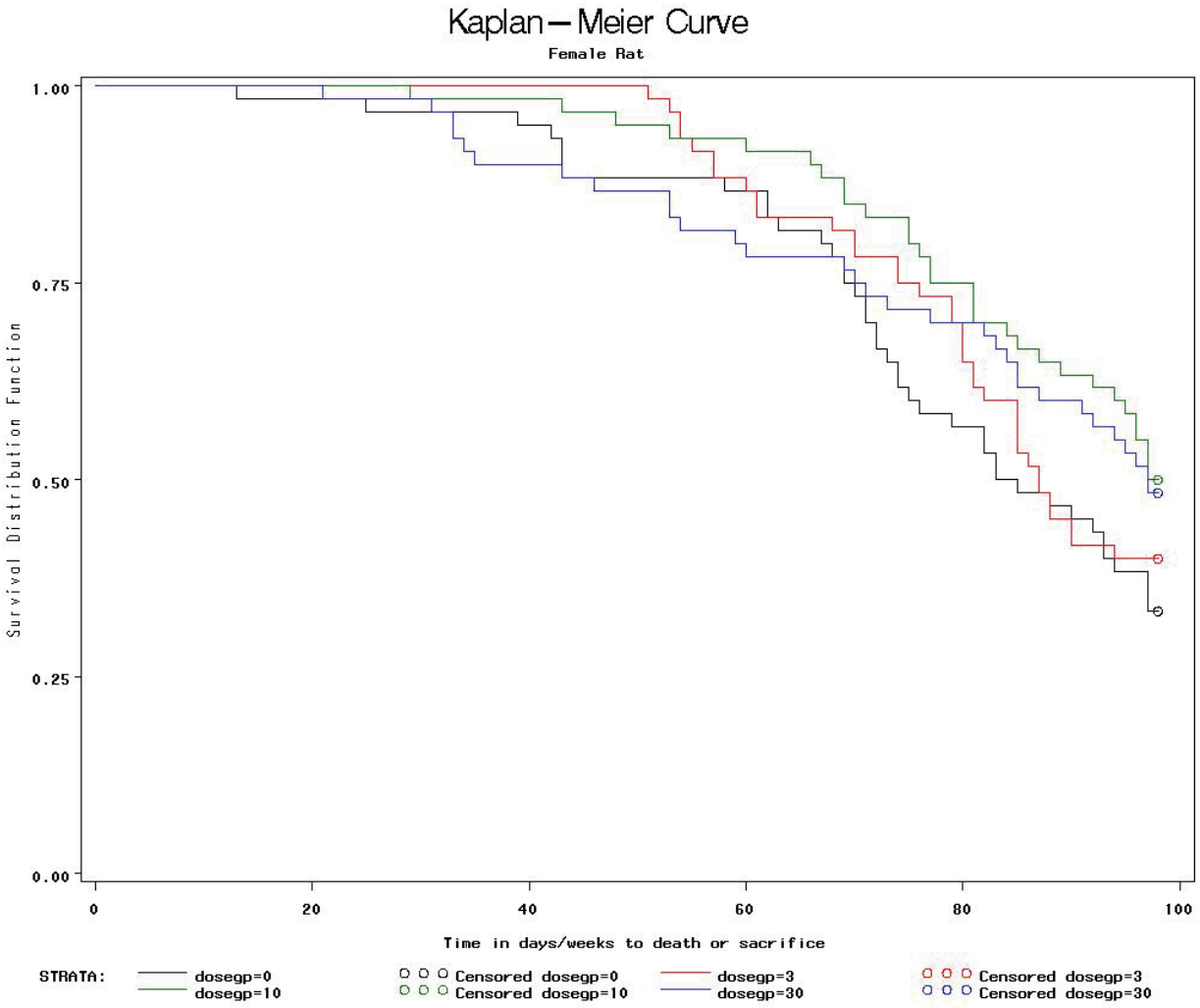
X-Axis: Weeks, Y-Axis: Survival rates

Figure 1A (2): Kaplan-Meier Survival Functions for Male Rats
(pair-fed control, low, medium and high dose groups)



X-Axis: Weeks, Y-Axis: Survival rates

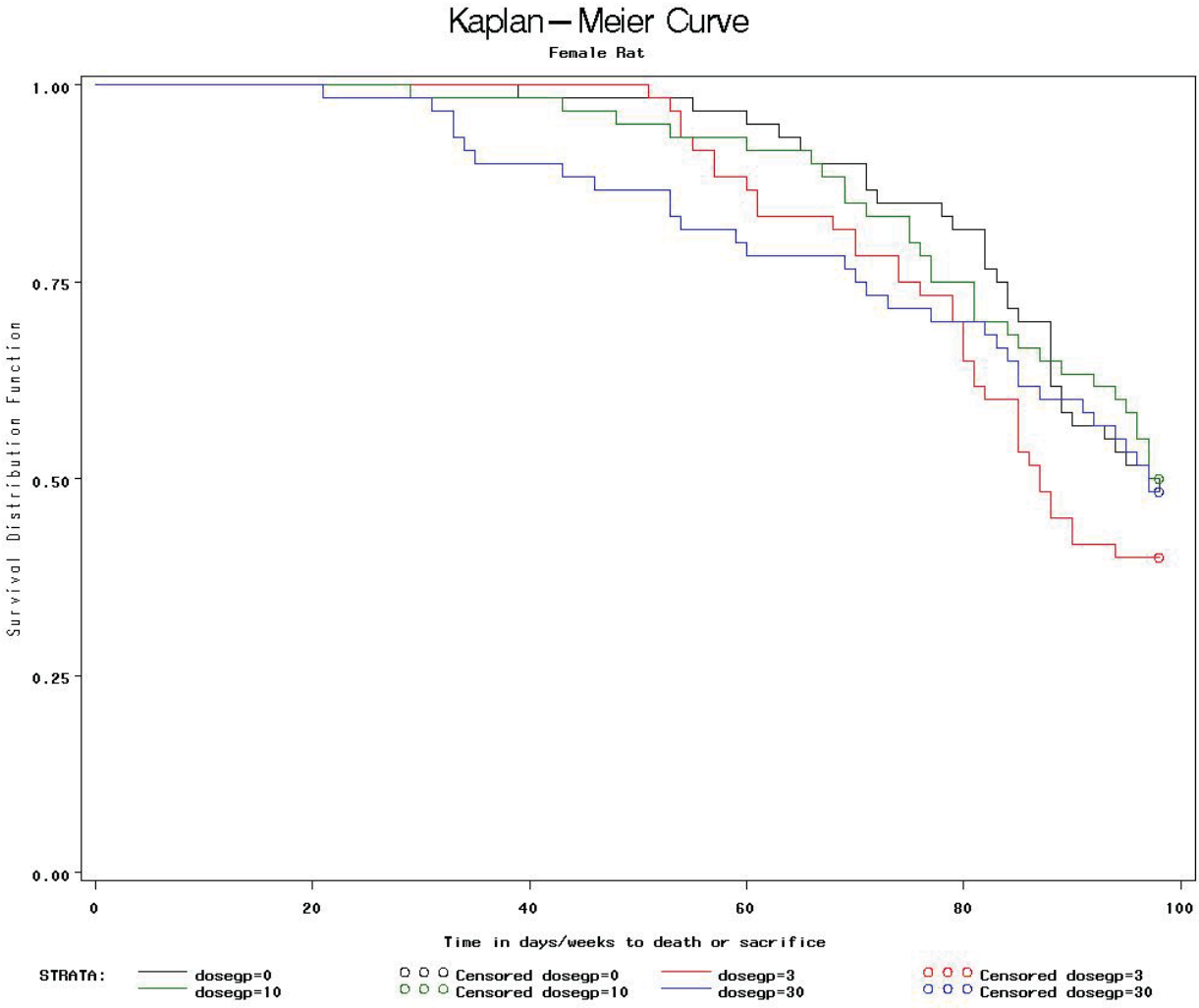
Figure 1B (1): Kaplan-Meier Survival Functions for Female Rats
(vehicle control, low, medium and high dose groups)



X-Axis: Weeks, Y-Axis: Survival rates

Figure 1B (2): Kaplan-Meier Survival Functions for Female Rats

(pair-fed control, low, medium and high dose groups)



X-Axis: Weeks, Y-Axis: Survival rates

5. References:

1. Bailer AJ, Portier CJ (1988). "Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples." *Biometrics*, 44, 417-431.

2. Bieler, G. S. and Williams, R. L. (1993). "Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity". *Biometrics* 49, 793-801.

3. Cox D. R. (1972) "Regression models and life tables", *Journal of the Royal Statistical Society*, B, 34, 187-220.

4. Gehan (1965) "A generalized Wilcoxon test for comparing arbitrarily singly censored samples", *Biometrika*, 52, 203-223.

5. Haseman, J (1983), "A re-examination of false-positive rates for carcinogenesis studies", *Fundamental and Applied Toxicology*, 3: 334-339.

6. Lin, K.K. and Rahman, M.A. (1998), "Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs", *Journal of Biopharmaceutical Statistics,* 8(1), 1-15.

7. Rahman, M.A. and Lin, K.K. (2008), "A comparison of False Positive Rates of Peto and Poly-3 Methods for Long-Term Carcinogenicity Data Analysis Using Multiple Comparison Adjustment Method Suggested by Lin and Rahman", *Journal of Biopharmaceutical Statistics,* 18:5, 849-858.

8. Peto, R., M.C. Pike, N.E. Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, Richards, and J.Wahrendorf (1980), "Guidelines for sample sensitive significance test for carcinogenic effects in long-term animal experiments", Long term and short term screening assays for carcinogens: A critical appraisal, International agency for research against cancer monographs, *Annex to supplement, World Health Organization, Geneva,* 311-426.

9. Tarone RE (1975), "Test for trend in life table analysis", *Biometrika*, 62: 679-82.

10. U.S. Department of Health and Human Services, "Guidance for Industry: Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals", Center for Drug E valuation and Research, Food and Drug Administration, Sliver Spring, Maryland, 2001.

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

/s/

MIN MIN
11/08/2010

KARL K LIN
11/09/2010
Concur with review

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

Statistical Review and Evaluation

CLINICAL STUDIES

NDA/Serial Number:	22-580/N-000
Drug Name:	QNEXA® Controlled Release Capsules (a combination of immediate-release phentermine hydrochloride beads and modified-release topiramate beads formulated for oral administration)
Indication(s):	Treatment of obesity, including weight loss and maintenance of weight loss, in conjunction with diet and exercise.
Applicant:	Vivus
Date(s):	Submitted December 28, 2009
Review Priority:	Standard
Biometrics Division:	Division of Biometrics 2 (HFD-715)
Statistical Reviewer:	Lee-Ping Pian, Ph.D.
Concurring Reviewers:	Todd Sahlroot, Ph.D., Deputy Director Tom Permutt, Ph.D., Director
Medical Division:	Division of Metabolic and Endocrine Products (HFD-510)
Clinical Team:	Mary Roberts, M.D., Medical Reviewer Eric Colman, M.D., Deputy Director
Project Manager:	Pat Madara

Table of Contents

1.1 Conclusions and Recommendations 3
1.2 Brief Overview of Clinical Studies 3

INTRODUCTION 4

2.1 Data Sources 4

3. STATISTICAL EVALUATION 4

3.1 Evaluation of Efficacy 4
3.2 Evaluation of Safety 20

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS 22

4.1 Gender, Race and Age 22

5. SUMMARY AND CONCLUSIONS 24

5.1 Statistical Issues and Collective Evidence 24
5.2 Conclusions and Recommendations 24

APPENDICES 25

1 Quality of Life 25
2. Boxplots by weight loss category and treatment group 28
3. Weight change in discontinuation cohort of patients 40
4. RBANS (Repeatable Battery for the Assessment of Neuropsychological Status) 43
5. On study patients who stopped treatment 50

EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Qnexa, a fixed-dose combination of phentermine hydrochloride (PHEN) and topiramate (TPM), was shown to be efficacious in reducing body weight for 2 of the 3 doses studied, PHEN/TPM 15/92 mg and PHEN/TPM 7.5/46 mg. The evidence for efficacy comes from factorial Study OB-301 which compared each combination dose to the respective components at the same dose and to placebo. This trial, by virtue of its design, was capable of providing evidence of efficacy in support of the combination drug rule which is the standard for evaluating efficacy in combination products. Results showed that all comparisons of the combinations to their components were statistically significant for the two co-primary efficacy endpoints, percentage change in weight from baseline and the proportion of patients with a minimum 5% weight loss from baseline. The combinations provided an additional 3 to 5 kg average reduction in body weight compared to weight changes seen in the monotherapy arms.

Studies OB-302 and OB-303 both compared two dose combinations to placebo. The combination doses tested in OB-302 and OB-303 were PHEN/TPM 15/92 and PHEN/TPM 3.75/23, and PHEN/TPM 15/92 and PHEN/TPM 7.5/46, respectively. All combinations provided statistically significant weight changes compared to placebo. However, these trials did not have monotherapy arms and therefore could not provide additional evidence of the efficacy of the combination doses in support of the combination drug rule. In particular, PHEN/TPM 3.75/23 mg was tested in study (OB-302) only, not Study OB-301. Therefore, the efficacy of 3.75/23 mg could not be confirmed from the data provided.

Discontinuations due to adverse events were dose related. Psychiatric and nervous system disorders resulted in the greatest number of discontinuations across all 3 studies. Only the 15/92 mg dose was significantly different from placebo in the frequency of discontinuations due to these disorders (Table 10, p. 20).

Based on our analysis of efficacy and selected safety data, we recommend that 7.5/46 mg PHEN/TPM should be the only indicated dose. However, the 3.75/23 mg dose may also be used in patients when 'considered for use as a treatment dose in some patients based on individual treatment goals' (sponsor's wording in proposed label).

1.2 Brief Overview of Clinical Studies

Phentermine hydrochloride up to 37.5 mg is indicated for short-term (a few weeks) weight reduction. Topiramate 200-400 mg/day (2 divided doses) is indicated for treatment of epilepsy and migraine headache prophylaxis.

Qnexa is a fixed combination drug product for weight loss. Qnexa is comprised of immediate-release phentermine hydrochloride beads (PHEN) and modified-release topiramate beads (TPM).

After randomization, weekly titration started with the lowest dose of the 4 proposed dose strengths PHEN/TPM, 3.75/23 mg. The three higher doses are PHEN/TPM 7.5/46 mg (mid

dose), PHEN/TPM 11.25/69 mg (three-quarter dose), and PHEN/TPM 15/92 mg (full dose). The PHEN/TPM 11.25/69 mg was intended only as a titration dose.

INTRODUCTION

2.1 Data Sources

The link below contains the study report and analysis dataset:

\\cdsesub1\EVSPROD\NDA022580\\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\obesity

The link below contains the updated dataset for RBAN:

\\cdsesub1\EVSPROD\NDA022580\\0035\m5\53-clin-stud-rep\537-crf-ipl\datasets\ob-301\analysis\adrbans.xpt

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

The phase 3 studies were all randomized, double-blind, multicenter, parallel-group, and placebo-controlled.

OB-301 was a factorial study which compared full-dose (PHEN/TPM 15/92 mg) and mid-dose (PHEN/TPM 7.5/46 mg) Qnexa with placebo and the respective PHEN and TPM components after 28 weeks of treatment.

The PHEN/TPM 3.75/23 mg dose was not studied against its components in study 301. The dose was only studied in OB-302 which compared full-dose (PHEN/TPM 15/92 mg) and low-dose (PHEN/TPM 3.75/23 mg) with placebo after 56 weeks of treatment. The treatment groups in OB-303 were PHEN/TPM 15/92 mg and PHEN/TPM 7.5/46 mg Qnexa and placebo.

Study Endpoints

The co-primary efficacy variables were percent weight loss at week 56 and percentage of patients with at least 5% weight loss at week 56.

Table 1 presents summary of the 3 studies.

Table 1 Phase 3 study summary

	OB-301	OB-302	OB-303

	OB-301	OB-302	OB-303
Study location (# sites)	USA (34)	USA (91)	USA (93)
dates	12/03/2007 – 9/30/2008	11/01/2007 – 5/19/2009	11/01/2007 – 6/30/2009
study duration	screening: 2 weeks titration: 4 weeks treatment: 24 weeks	screening: 2 weeks titration: 4 weeks treatment: 52 weeks	screening: 2 weeks titration: 4 weeks treatment: 52 weeks
study population	adults ≤ 70 years of age BMI: ≥ 30 kg/m^2 & ≤ 45kg/m^2 No diabetics	adults ≤ 70 years of age BMI: ≥ 35 kg/m^2 No diabetics	adults ≤70 years of age BMI ≥ 27 kg/m^2 & ≤ 45kg/m^2 ≥ 2 obesity-related co-morbid conditions
Treatment groups	total n=756 1:1:1:1:1:1:1 stratified by gender Placebo (n=109) PHEN 7.5 mg (109) PHEN 15 mg (108) TPM 46 mg (108) TPM 92 mg (107) PHEN/TPM 7.5/46 mg (107) PHEN/TPM 15/92 mg (108)	n=1267 2:1:2 stratified by gender (male ≥ 20%) Placebo (514) PHEN/TPM 3.75/23 mg (241) PHEN/TPM 15/92 mg (512)	n=2487 2:1:2 stratified by gender (male ≥ 20%) and diabetes status Placebo (994) PHEN/TPM 7.5/46 mg (498) PHEN/TPM 15/92 mg (995)
Study objective	Combination superior to placebo and components	Combination superior to placebo	Combination superior to placebo
co-primary endpoints	Percent weight loss at Week 28 with LOCF; and Percentage of patients with ≥ 5% weight loss at week 28 with LOCF	Percent weight loss at Week 56; and Percentage of patients with at least 5% weight loss at Week 56	Percent weight loss at Week 56; and Percentage of patients with at least 5% weight loss at Week 56
Secondary endpoints	Percentage of patients with ≥ 10% weight loss at Week 28 with LOCF, Change in waist circumference from baseline to Week 28 with LOCF Changes in IWQOL*-Lite composite and individual domain scores at Week 28 with LOCF	Absolute weight loss at Week 56 Percentage of patients with ≥ 1**0%** weight loss at Week 56 and Change in waist circumference from baseline to Week 56	Absolute weight loss at Week 56 Percentage of patients with ≥ 10% weight loss at Week 56 and Change in waist circumference from baseline to Week 56

*IWQOL=Impact of Weight on Quality of Life

Patient Disposition, Demographic and Baseline Characteristics

The analysis of patient disposition was based on the set of randomized patients (100%). The safety population was defined as all randomized patients who received at least one dose of study drug. The sponsor's ITT population was defined as all randomized patients who took at least one dose of the study drug, had a baseline and at least one post-baseline measurement of body weight (regardless of receiving study drug or being off study drug). The modified ITT population (MITT) was defined as all ITT patients who had at least one post baseline measurement of body

weight within 7 days of the last dose of study drug. The MITT was smaller than the ITT population by approximately 2% for all 3 studies. Tables 2-4 summarize the disposition of patients for the 3 studies, respectively.

Approximately 66% of patients in Study 301 completed the 28-week study on drug. The 56-week completion rates for studies 302 and 303 were 54% and 62%, respectively.

Table 2 Patient disposition – Study 301

n(%)	Placebo (N=109)	PHEN 7.5 (N=109)	TPM 46 (N=108)	PHTN/TPM 7.5/46 (N=107)	PHEN 15 (N=108)	TPM 92 (N=107)	PHTN/TPM 15/92 (N=108)	Total (N=756)
Safety	109	109	106 (98)	106 (99)	108	107	108	753
ITT	103 (95)	104 (95)	102 (94)	103 (96)	106 (98)	105 (98)	103 (95)	726 (96)
MITT*	102 (94)	100 (92)	100 (93)	98 (92)	105 (97)	104 (97)	101 (94)	710 (94)
Completed visits	74 (68)	79 (73)	78 (72)	78 (73)	80 (74)	77 (72)	75 (69)	541 (72)
Completed*	69 (63)	74 (68)	72 (67)	73 (68)	72 (67)	67 (63)	68 (63)	495 (66)
Discontinued	40 (37)	35 (32)	36 (33)	34 (32)	36 (33)	40 (37)	40 (37)	261 (35)
AE	8 (7)	10 (9)	8 (7)	16 (15)	11 (10)	18 (17)	23 (21)	94 (12)
Lost follow-up	12 (11)	13 (12)	11 (10)	6 (6)	7 (7)	8 (9)	9 (8)	66 (9)
Consent withdrew	9 (8)	7 (6)	6 (6)	4 (4)	8 (7)	5 (5)	3 (3)	42 (6)

*receiving study drug

Table 3 Patient disposition – Study 302

n(%)	Placebo	PHEN/TPM PHEN/TPM 3.75/23	PHEN/TPM PHEN/TPM 15/92	Total
Randomized	514	241	512	1267
Safety	513	240	511	1264
ITT	498 (97)	234 (97)	498 (97)	1230 (97)
MITT	485 (94)	229 (95)	487 (95)	1201 (95)
Completed all study visits	272 (53)	147 (61)	340 (66)	759 (60)
Completed all visits on study drug	241 (47)	138 (57)	301 (59)	680 (54)
Discontinued study drug	272 (53)	102 (42)	210 (41)	584 (46)
Adverse event	43 (8)	28 (12)	83 (16)	154 (12)
Subject lost to follow-up	89 (17)	27 (11)	53 (10)	169 (13)
Subject withdrew consent	86 (17)	28 (12)	39 (8)	153 (12)
Lack of efficacy	23 (4.5)	6 (2.5)	6 (1.2)	35 (2.8)

Table 4 Patient disposition – Study 303

n(%)	Placebo	PHTN/TPM 7.5/46	PHTN/TPM 15/92	Total
Randomized	994	498	995	2487
Safety Set	993	498	994	2485
Intent-to-Treat Set	979 (99)	488 (98)	981 (99)	2448 (98)
Modified Intent-to-Treat Set	957 (96)	482 (97)	963 (97)	2402 (97)
Completed all study visits	616 (62)	374 (75)	733 (74)	1723 (69)
Completed all visits on study drug	564 (57)	344 (69)	634 (64)	1542 (62)
Discontinued study drug	429 (43)	154 (31)	360 (36)	943 (38)
Adverse event	89 (9)	58 (12)	192 (19)	339 (14)
Subject withdrew consent	139 (14)	34 (7)	69 (7)	242 (10)
Subject lost to follow-up	126 (13)	41 (8)	62 (6)	229 (9)
Lack of efficacy	39 (3.9)	3 (0.6)	5 (0.5)	47 (1.9)

Figure 1 displays Kaplan Meire curves for time to study drug discontinuation by treatment groups in the randomized population. Figure 2 displays time to discontinuation of study drug if the reason for discontinuation was due to adverse events.

Figure 1 Time to discontinuation of study drug

RAND
1

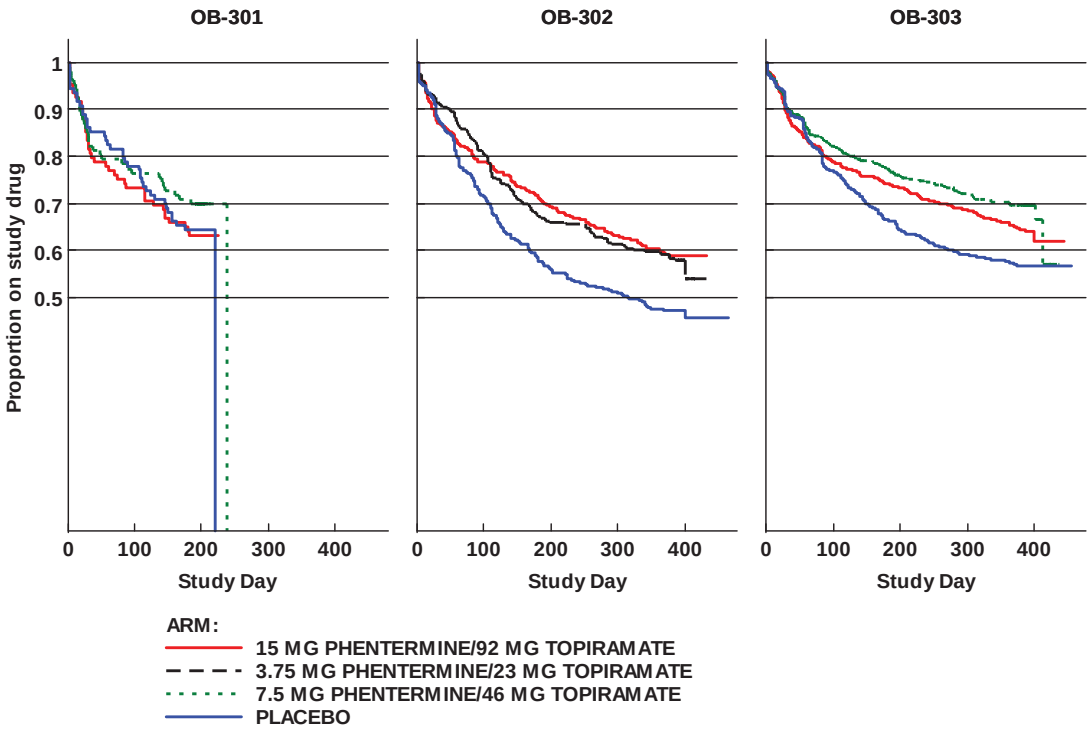

OB-301
OB-302
OB-303
Proportion on study drug
1
0.9
0.8
0.7
0.6
0.5
0
100
200
300
400
Study Day
ARM:
15 MG PHENTERMINE/92 MG TOPIRAMATE
3.75 MG PHENTERMINE/23 MG TOPIRAMATE
7.5 MG PHENTERMINE/46 MG TOPIRAMATE
PLACEBO

Figure 2 Time to discontinuation of study drug for any AE

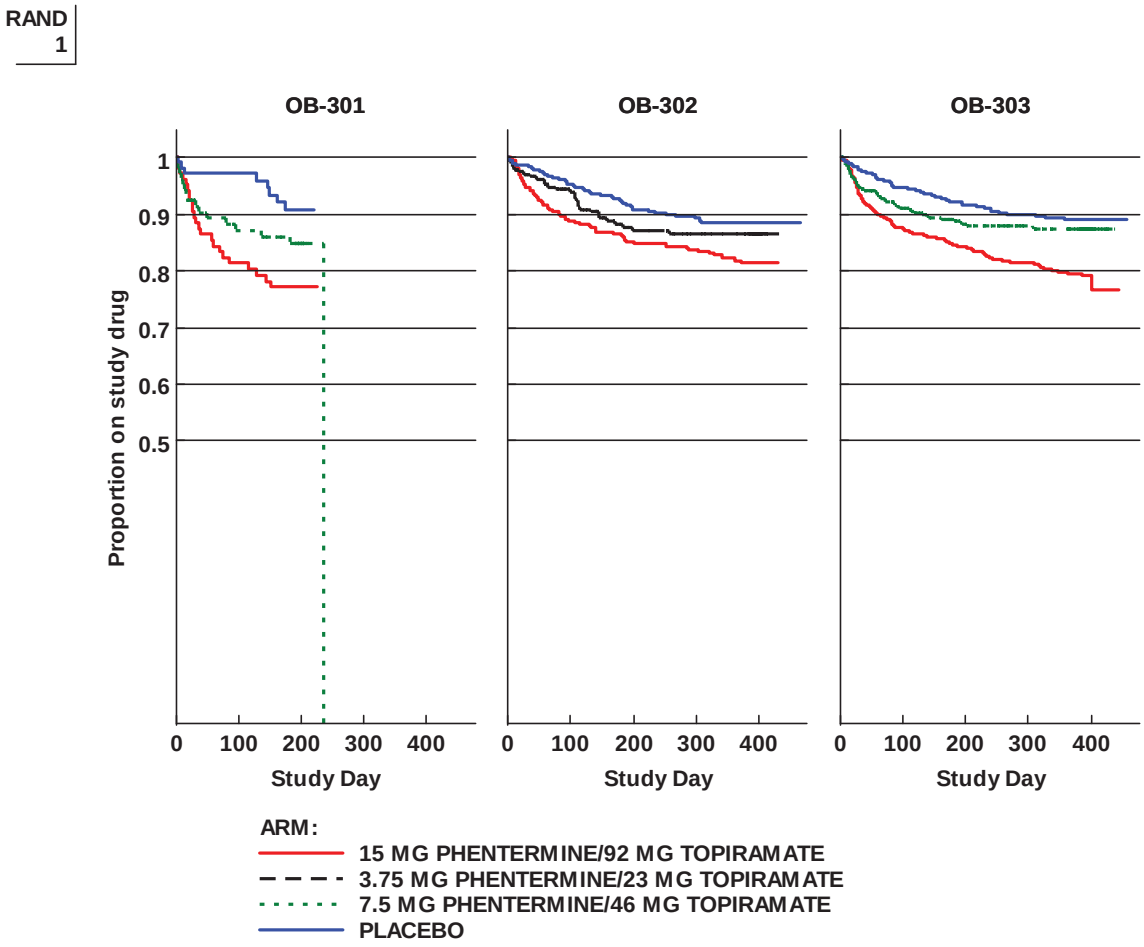


Table 5 summarizes baseline and demographic characteristics by study for all randomized patients. Mean age were 46, 43, and 51 years for studies 301, 302 and 303, respectively. The majority of patients were females and Caucasian. Average weight was a little over 100 kg for studies 301 and 303 and 116 kg for study 302. The BMI were 36 kg/m^2, 37 kg/m^2 for studies 301 and 303 and 42 kg/m^2 for study 302 (inclusion criteria BMI≥ 35 kg/m^2 for study 302 and BMI≥ 30 kg/m^2 and BMI≤ 45 kg/m^2 for studies 301 and 303).

Table 5 Demographics and baseline characteristics by study*

	Study 301 n=756	Study 302 n=1267	Study 303 n=2487
Age (years)			
Mean (SD) [min, max]	46 (11.9) [18, 71]	43 (11.8) [18, 70]	51 (10.4) [19, 71]
Gender n (%)			
Female	599 (79%)	1050 (83%)	1737 (70%)
Race %			

	Study 301 n=756	Study 302 n=1267	Study 303 n=2487
Caucasian	79%	79%	86%
African	19%	18%	12%
Other	2%	3%	2%
Weight (kg)			
Mean (SD) [min, max]	101 (15.5) [65, 167]	116 (21.2) [71, 217]	103 (17.9) [58, 163]
Height (cm)			
Mean (SD) [min, max]	167 (8.6) [145, 199]	166 (8.8) [132, 201]	168 (9.7) [127, 201]
Body mass index (kg/m^2)			
Mean (SD) [min, max]	36 (4.1) [30, 45]	42 (6.2) [34, 79]	37 (4.5) [21, 51]
Waist circumference (cm)			
Mean (SD) [min, max]	111 (11.1) [85, 157]	121 (14.4) [88, 198]	113 (12.3) [80, 157]
SBP (mmHg)			
Mean (SD) [min, max]	122 (13) [82, 163]	122 (11) [84, 166]	128 (14) [73, 188]
DBP (mmHg)			
Mean (SD) [min, max]	79 (9) [26, 100]	77 (8) [51, 100]	81 (9) [50, 120]
HR (bpm)			
Mean (SD) [min, max]	73 (10) [45, 114]	73 (9) [46, 108]	72 (10) [40, 111]

Other baseline characteristics by study

	Study 301 (n=726)	**Study 302 (n=1267)**	**Study 303 (n=2487)**
LDL cholesterol (mg/dL)			
Mean (SD) [min, max]	126 (32) [46, 291]	121 (31) [30, 271]	123 (35) [8, 292]
HDL cholesterol (mg/dL)			
Mean (SD) [min, max]	52 (15) [17, 138]	50 (12) [23, 112]	49 (14) [8, 138]
Total cholesterol (mg/dL)			
Mean (SD) [min, max]	205 (36) [110, 405]	194 (35) [85, 363]	204 (40) [78, 395]
Triglycerides (mg/dL)			
Mean (SD) [min, max]	135 (65) [31, 467]	116 (39) [33,262]	162 (74) [33, 656]
Fasting serum glucose (mg/dL)			
Mean (SD) [min, max]	94 (10) [62, 125]	93 (9) [42, 141]	106 (22) [43, 295]
HbA_{1c}			2478
Mean (SD) [min, max]	5.5 (0.4) [4, 6.9]		5.9 (0.8) [4.1, 11.9]

*n might vary by baseline characteristics

Statistical Methodologies

In study OB-301, the confirmation of efficacy for the combination was based on a set of three pair-wise comparisons (i.e., combination versus each components and combination versus placebo). All three pairwise comparisons must reach the 5% significance level for both of the co-primary endpoints in order for that dose to be considered effective according to the "combination rule".

The multiple comparison issue posed by two combination doses was addressed by a step-down procedure which tested the high dose PHEN/TPM 15/92 mg first and, if significant, then tested the second combination dose.

The primary objective of Studies 302 and 303 was to demonstrate the superiority of the studied combination to placebo on the co-primary efficacy endpoints. These studies did not have monotherapy arms using doses corresponding to the doses in the combination drug arms. Therefore, due to the limitations in the study designs, these trials cannot address the efficacy requirements of the combination rule which is necessary for the efficacy evaluation of combination products.

Results and Conclusions

Body weight percent change from baseline:

The sponsor's primary efficacy analysis used the ITT population which included both on-drug and off-drug patients. This reviewer used the MITT population which included data only from patients who were on study drug at the time of measurement. The percent weight changes in the 2 populations were very similar with approximately 0.3% more reduction in the on drug patients (MITT). Tables 6-8 display the analysis of covariance results for the 3 respective studies. Figures 3 and 4 display the cumulative distribution and box plot by treatment group for % weight change from baseline to week 56 in the MITT (LOCF) population. Figures 5 and 6 display the LSM (least squared mean) between treatment difference and the 95% confidence interval (CI). Figures 7 and 8 display the percent weight change over time by treatment group for the completers.

Table 6 ANCOVA results of mean percent change from baseline – MITT LOCF Week 28 – Study 301

LSM at baseline and % Wt change from baseline				Between treatment difference in % Wt change from baseline		
Trt	**n**	**Baseline (kg) (SE)**	**% Change (SE)**	**combo vs. placebo or component**	**LSM difference (CI)**	**p-value**
15P/ 92T	101	104 (1.4)	-9.5 (0.6)	15P/92T vs. Plb	**-7.8 [-9.4, -6.1]**	<0.01
15P	105	107 (1.4)	-6.1 (0.6)	15P/92T vs. 15P	**-3.5 [-5.1, -1.9]**	<0.01
92T	104	110 (.14)	-6.6 (0.6)	15P/92T vs. 92T	**-2.9 [-4.6, -1.3]**	<0.01
7.5P/ 46T	98	108 (1.4)	-8.9 (0.6)	7.5P/46T vs. Plb	**-7.1 [-8.8, -5.5]**	<0.01
7.5P	100	106 (1.4)	-5.7 (0.6)	7.5P/46T vs. 7.5P	**-3.2 [-4.9, -1.6]**	<0.01
46T	100	105 (1.4)	-5.3 (0.6)	7.5P/46T vs. 46T	**-3.7 [-5.3, -2.0]**	<0.01
Plb	102	105 (1.4)	-1.8 (0.6)			

*ANCOVA model: fixed effects of treatment and gender and baseline weight as a covariate

Table 7 ANCOVA results of mean percent change from baseline – MITT, LOCF Week 56 – Study 302

Treatment	n	LSM	SE	Treatment difference	LSM (SE) 95% CI	p
15 PHEN/92 TPM	487	-11.5	(0.4)	PHEN/TPM 15/92 vs. Placebo	-9.9 (0.5) [-10.8, -9.0]	<0.001
3.75 PHEN/23 TPM	229	-5.3	(0.5)	PHEN/TPM 3.75/23 vs. Placebo	-3.8 (0.6) [-4.9, -2.6]	<0.001
PLACEBO	485	-1.6	(0.4)	PHEN/TPM 15/92 vs. PHEN/TPM 3.75/23	-6.1 (0.6) [-7.3, -5.0]	<0.001

Table 8 ANCOVA results of mean percent change from baseline – MITT, LOCF Week 56 – Study 303

Treatment	n	LSM	SE	Treatment difference	LSM (SE) 95% CI	p
15 PHEN/92 TPM	963	-10.4	(0.2)	PHEN/TPM 15/92 vs. Placebo	-8.6 (0.3) [-9.3, -7.9]	<0.001
7.5 PHEN/46 TPM	482	-8.3	(0.3)	PHEN/TPM 7.5/46 vs. Placebo	-6.5 (0.4) [-7.3, -5.6]	<0.001
PLACEBO	957	-1.8	(0.2)	PHEN/TPM 15/92 vs. PHEN/TPM 7.5/46	-2.1 (0.4) [-2.9, -1.3]	<0.001

Figure 3 Cumulative distribution and box plot of % weight loss – MITT, LOCF

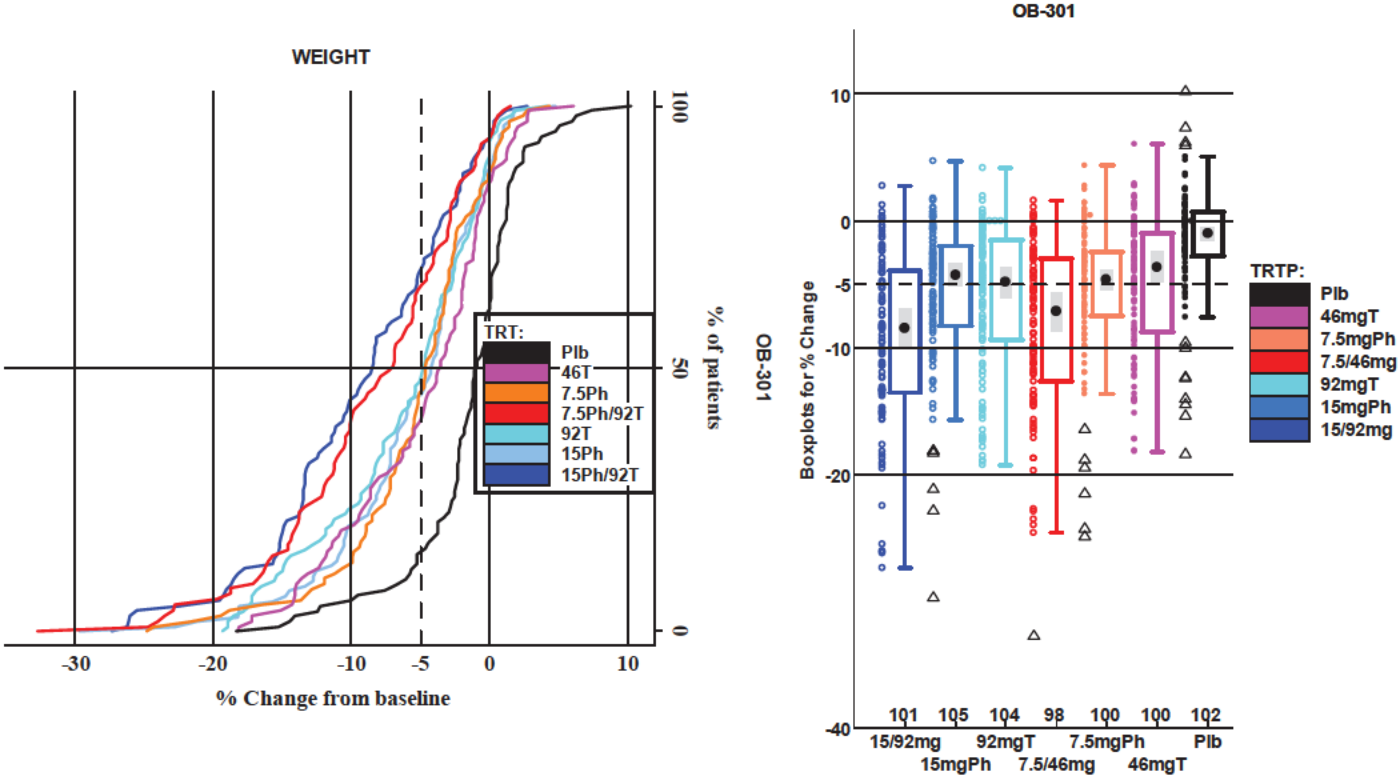


Figure 4 Box plot and cumulative distribution of % weight change from baseline to week 56
Studies 302, 303 – MITT, LOCF

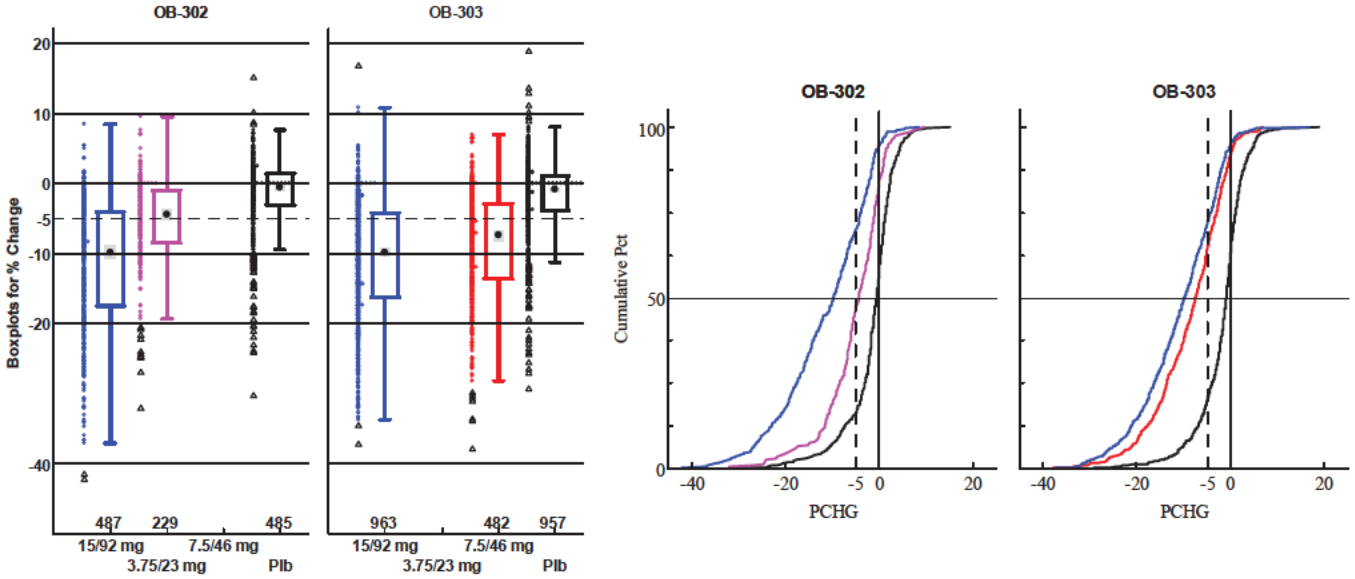


Figure 5 LSM treatment differences of combinations vs. components and placebo in % Weight change from baseline to Week 28 – MITT (patients on drug), LOCF Study 301

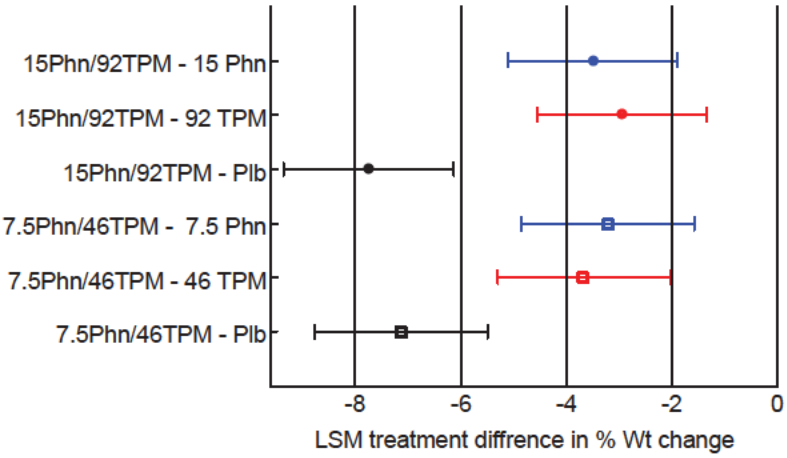


Figure 6 LSM treatment differences in % Weight change from baseline to Week 56 – MITT (patients on study drug), LOCF

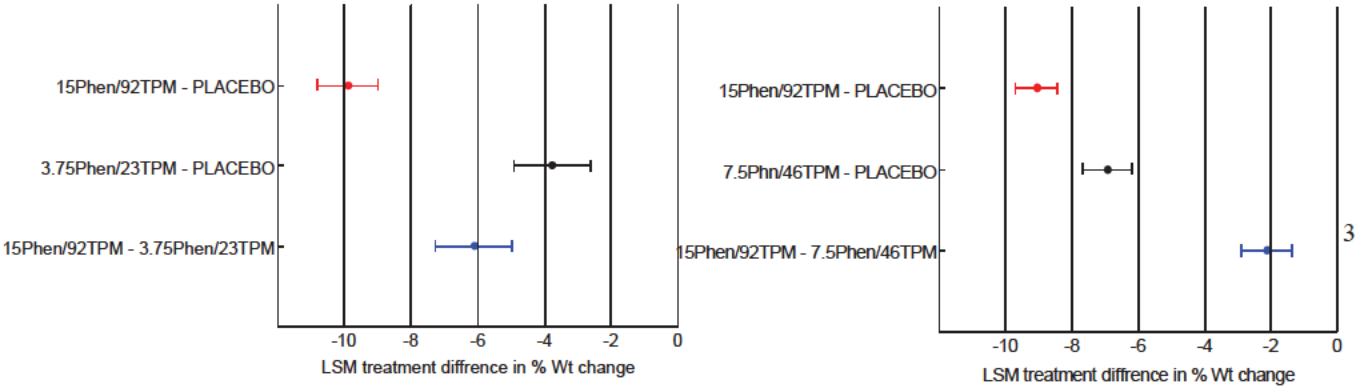
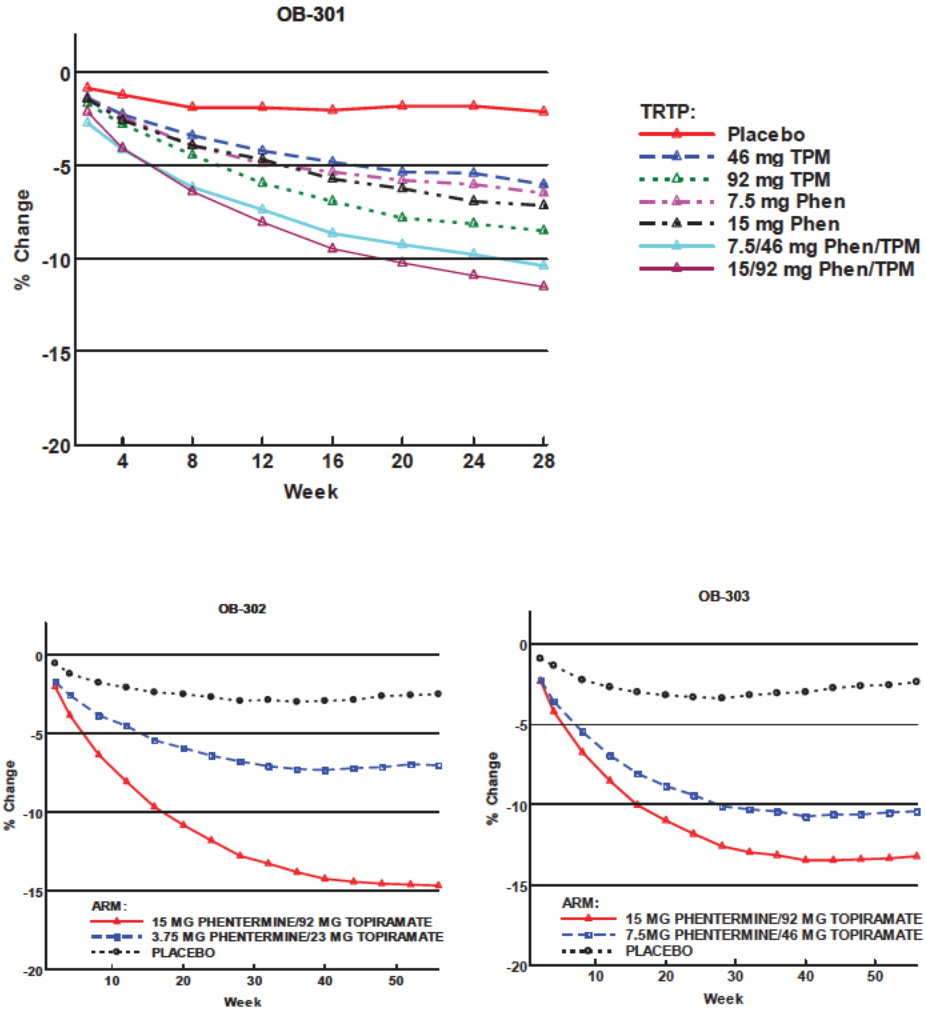


Figure 7 Mean % weight change from baseline by time and treatment group – Patients who completed all visits on study drug



Percentage of patients with ≥ 5% weight loss from baseline (Co-primary endpoint)

Table 9 displays statistical analyses of the percentage of patients with ≥ 5% weight loss from baseline. All comparisons between the combinations and placebo were statistically significant ($p<0.01$) as well as the combinations vs. the respective components.

Table 9 Risk differences compared to placebo [95% CI] for the percentage of patients with ≥ 5% weight loss from baseline

% of responders	placebo	PHEN/TPM 3.75/23 mg	PHEN/TPM 7.5/46 mg	PHEN/TPM 15/92 mg
301	16%		62%	66%
302	16%	48%		70%
303	20%		64%	72%
vs. placebo				
301			**47% [35, 58]**	**53% [49, 57]**
302		**31% [24, 39]**		**54% [48, 59]**
303			**44% [40, 49]**	**52% [48, 56]**

Combination Study 301	PHEN/TPM 7.5/46 mg	PHEN 7.5 mg	TPM 46 mg	PHEN/TPM 15/96 mg	PHEN 15 mg	TPM 92 mg
% of responders	62%	43%	39%	66%	46%	49%
vs. component		**19% [6, 32]**	**23% [10, 36]**		**20% [7, 33]**	**17% [4, 31]**

Figure 8 Percentage of patients with ≥ 5% weight loss at Week 28 – MITT, LOCF Study 301

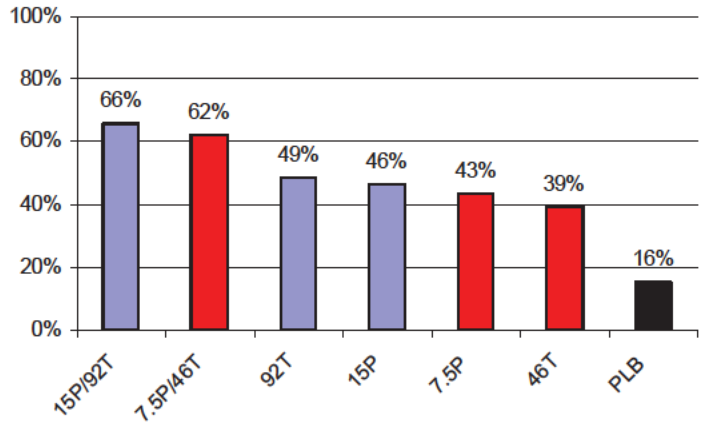


Figure 9 Percentage of patients with ≥ 5% weight loss at Week 56 – MITT, LOCF Study 302

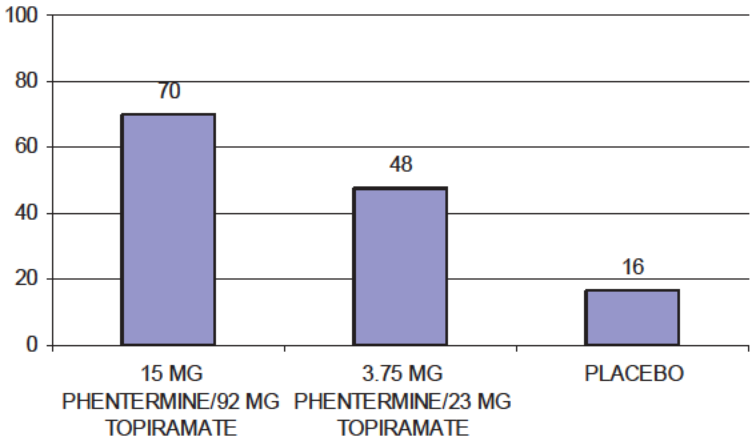
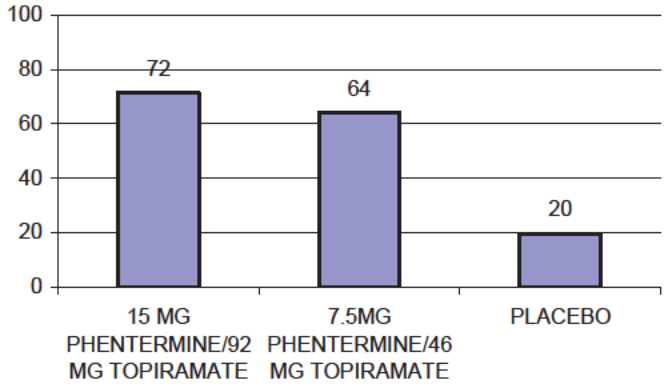


Figure 10 Percentage of patients with ≥ 5% weight loss at Week 56 – MITT, LOCF Study 303



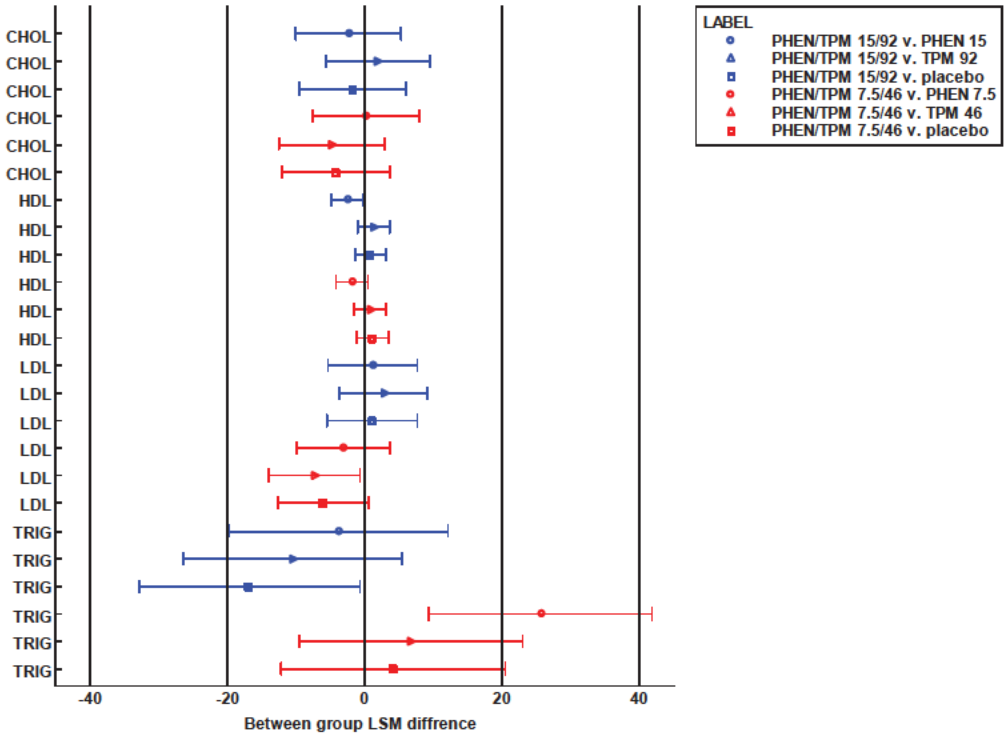
Other endpoints

Figures 11-14 display treatment differences between the combinations and components or placebo in change from baseline to week 28 (MITT, LOCF) for Study 301. Other than absolute changes in weight and waist circumference, none of the endpoints was significantly different from both of the respective components and the placebo for both combination doses, PHEN/TPM 15/92 mg and 7.5/32 mg. Figure 15 displays the treatment differences between the combinations and placebo for Studies 302 and 303. The components were not studied in the 2 trials.

Lipids

None of the combinations was statistically significantly different from placebo and both of the components. The triglycerides change from baseline was significantly worse in the PHEN/TPM 7.5/46 mg treatment group compared to the PHEN 7.5 mg treatment alone.

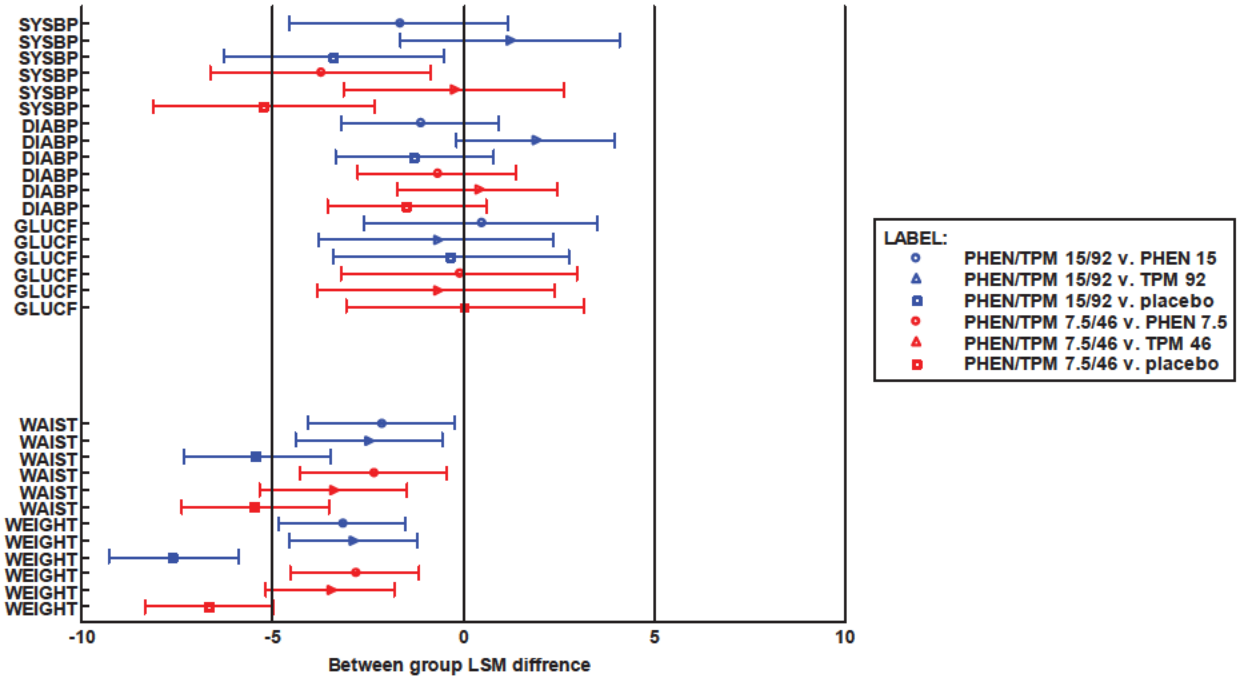
Figure 11 Between-group least squared mean difference (95% CI) – Study 301



SBP, DBP, FBG, Waist and Weight

Figure 12 shows only the absolute changes in weight and waist circumference were significantly different from their respective components and placebo for both combinations.

Figure 12 Between-group least squared mean difference (95% CI) – Study 301



HbA1c

For the PHEN/TPM 7.5/46 mg group, the significant between group differences in HbA1c change from baseline were approximately -0.1% or less compared to placebo and the components. The high dose combination was not significantly better than the TPM 92 mg component.

Figure 13 Between-group least squared mean difference (95% CI) – Study 301

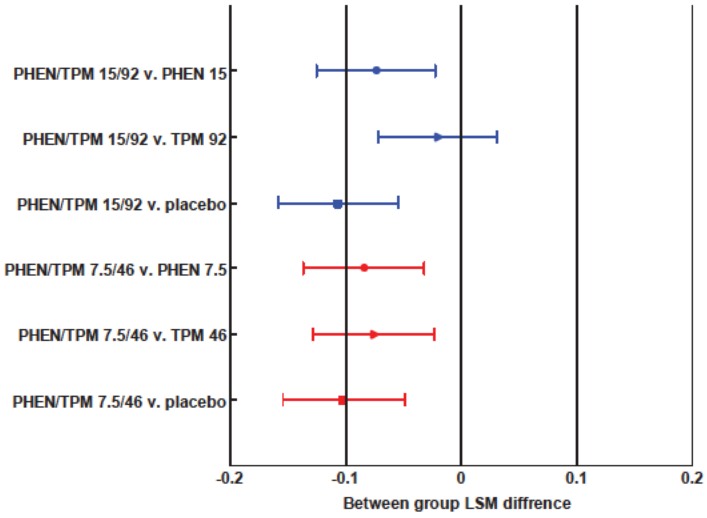
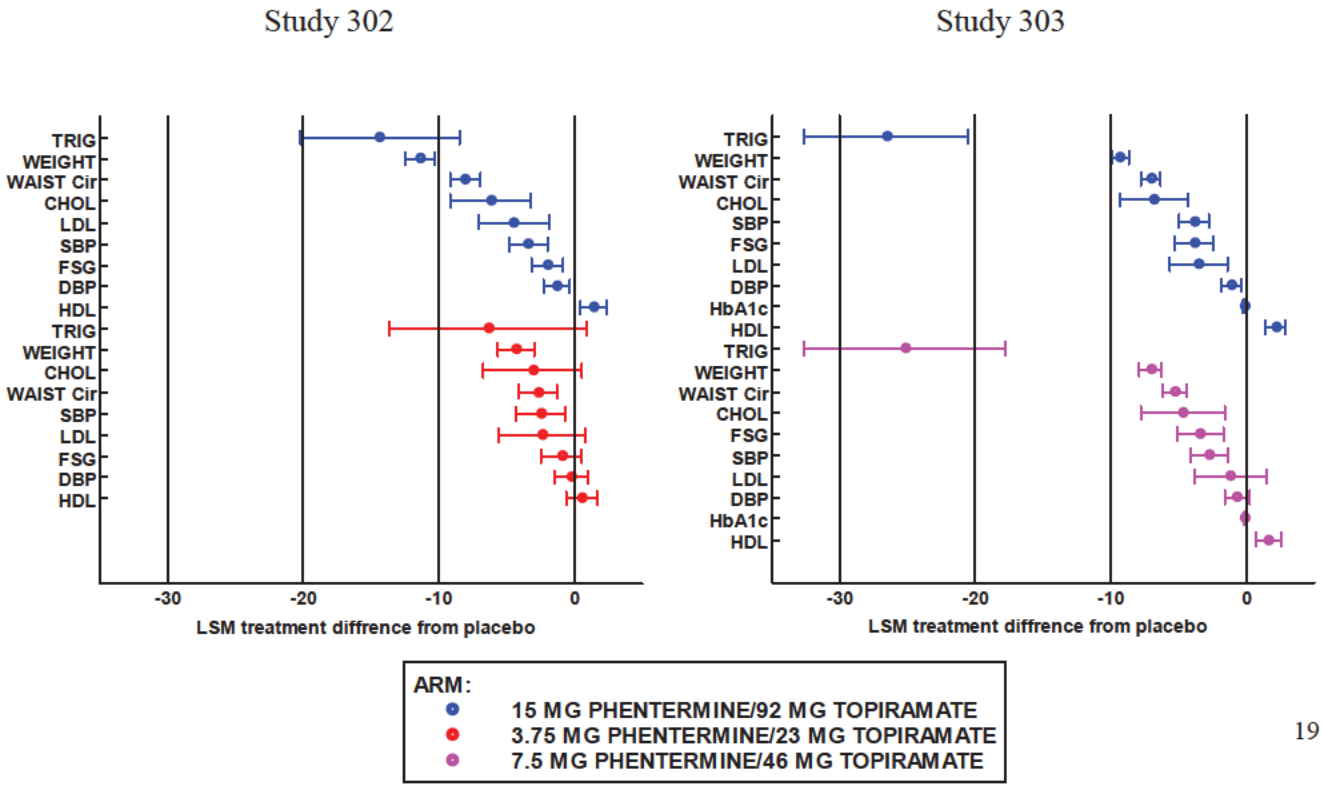


Figure 14 LSM difference (95% CI) from placebo



3.2 Evaluation of Safety

The 2 most frequent adverse events prompting discontinuation of study medication were psychiatric and nervous system disorders, both of which were dose related.

The Cochran Armitage trend test for dose response stratified by study was significant ($p<0.001$) for both types of adverse events. Table 6 displays the common odds ratio (95% CI) and p-value for each dose versus placebo. The PHEN/TPM 15/92 mg dose was significantly worse than placebo in AE discontinuation from the study drug. The other 2 doses were not significantly different from placebo which applies to both AEs (Table 10). Figures 15 and 16 display the Kaplan Meier curves for the AE discontinuation from the study drug for the psychiatric and neurological disorder, respectively.

Table 10 Analyses of discontinuation of study drug due to psychiatric and nervous system AE – all randomized patients

	Placebo	PHEN/TPM 3.75/23	PHEN/TPM 7.5/46	PHEN/TPM 15/92
301	n=109		n=107	n=108
302	n=514	n=241		n=512
303	n=994		n=498	n=995
Psychiatric				
301	1 (0.9%)		2 (1.9%)	6 (5.6%)
302	4 (0.8%)	6 (2.5%)		22 (4.3%)
303	14 (1.4%)		11 (2.2%)	46 (4.6%)
Stratified by study				
OR vs. placebo [95% CI], 2-sided p-value		**3.3 [0.8, 15.8] p=0.08**	**1.6 [0.7, 3.7] p=0.23**	**4 [2.4, 7.1] p<0.001**
Nervous system				
301	1 (0.9%)		3 (2.8%)	9 (8.3%)
302	10 (2%)	3 (1.2%)		15 (2.9%)
303	13 (1.3%)		10 (2%)	44 (4.4%)
stratified by study				
OR vs. placebo [95% CI], 2-sided p-value		**0.6 [0.1, 2.5] p=0.57**	**1.7 [0.8, 3.9] p=0.22**	**2.9 [1.8, 4.9] p<0.001**

OR=odds ratio

Figure 15 Time to discontinuation of study drug due to psychiatric disorder

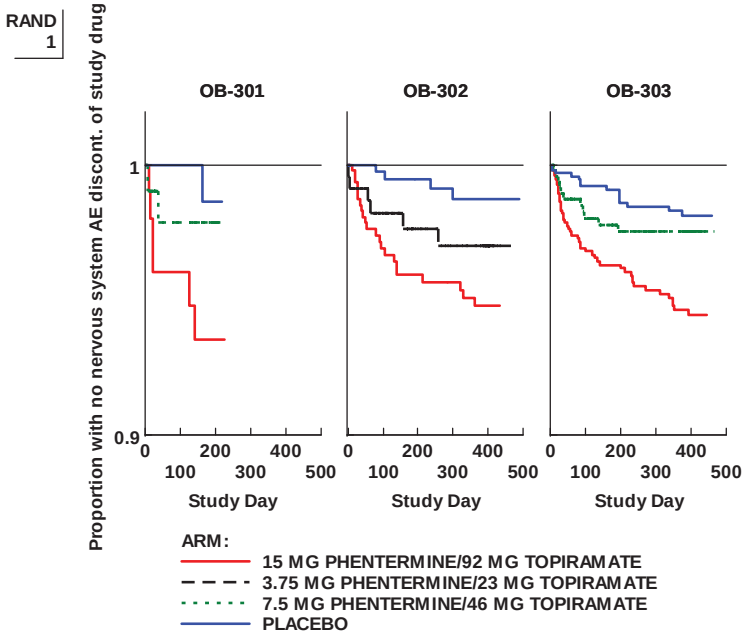
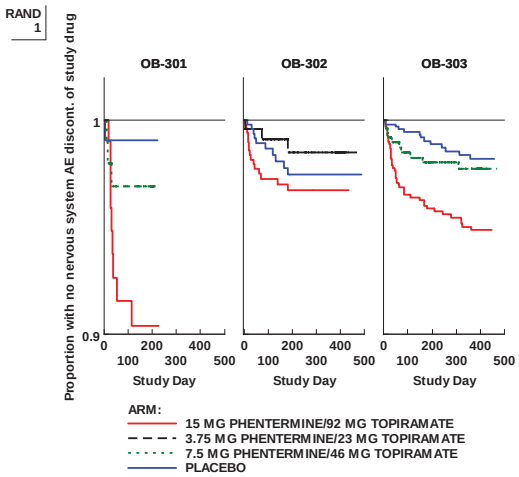


Figure 16 Time to discontinuation of study drug due to nervous system AE



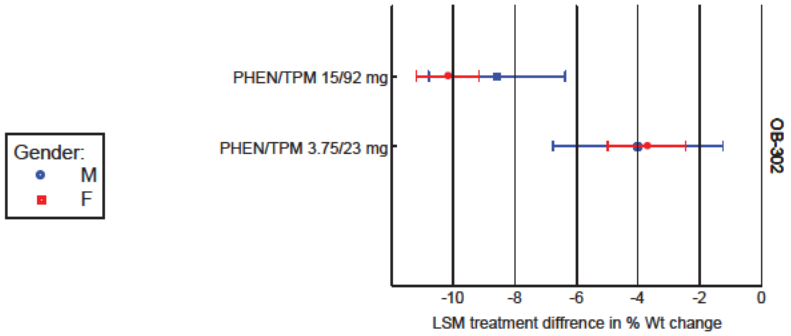
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

Treatment-by-gender interaction was not significant for Study 301 (p=0.4) at week 28 and Study 302 (p=0.3) at week 56.

Study 302

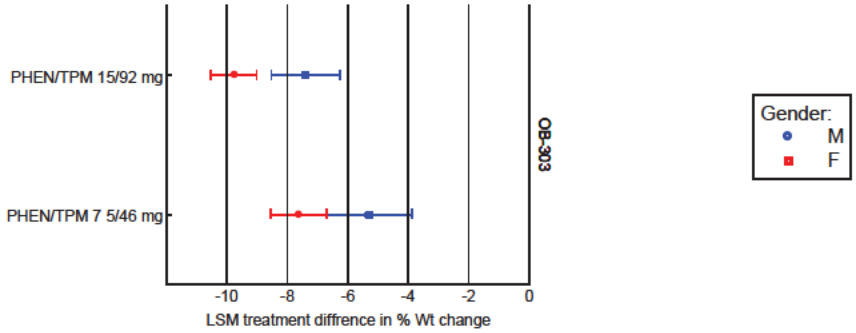
Figure 17 LSM difference between treatment and placebo by gender in % weight change– Study 302 MITT, LOCF



Study 303

Treatment-by-gender interaction was significant for both treatment groups (p<0.01). The treatment differences between test drug and placebo in weight percent change from baseline were -9.8% and -7.4 %, respectively for females and males (15 mg PHEN/92 mg TPM) and -7.6 % and -5.3 %, respectively (7.5 mg PHEN/46 mg TPM).

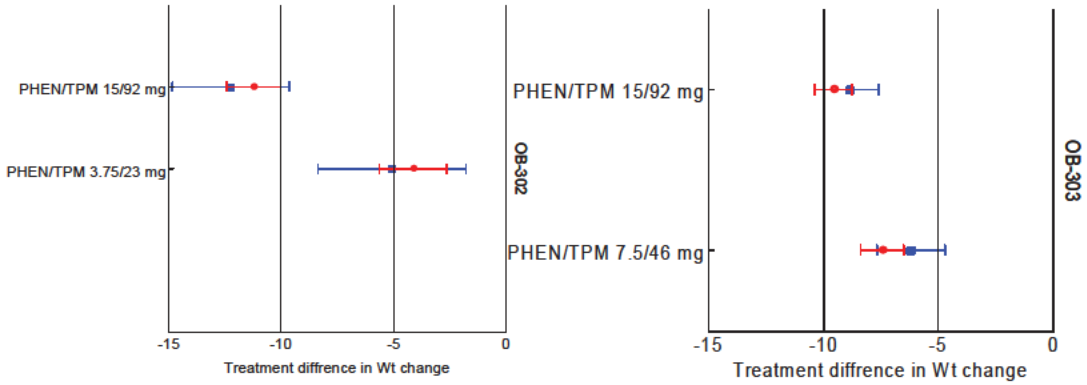
Figure 18 LSM difference between treatment and placebo by gender in % weight change– Study 303 MITT, LOCF



The treatment-by-gender interaction is quantitative (in size) and not qualitative (in direction). On average, females lost 2% more than males in treatment groups PHEN/TPM 15/92 mg and PHEN/TPM 7.5/46 mg in Study 303.

The absolute body weight changes from baseline were similar in males and females with no treatment-by-gender interaction.

Figure 19 LSM difference between treatment and placebo by gender in absolute weight change– Study 303 MITT, LOCF

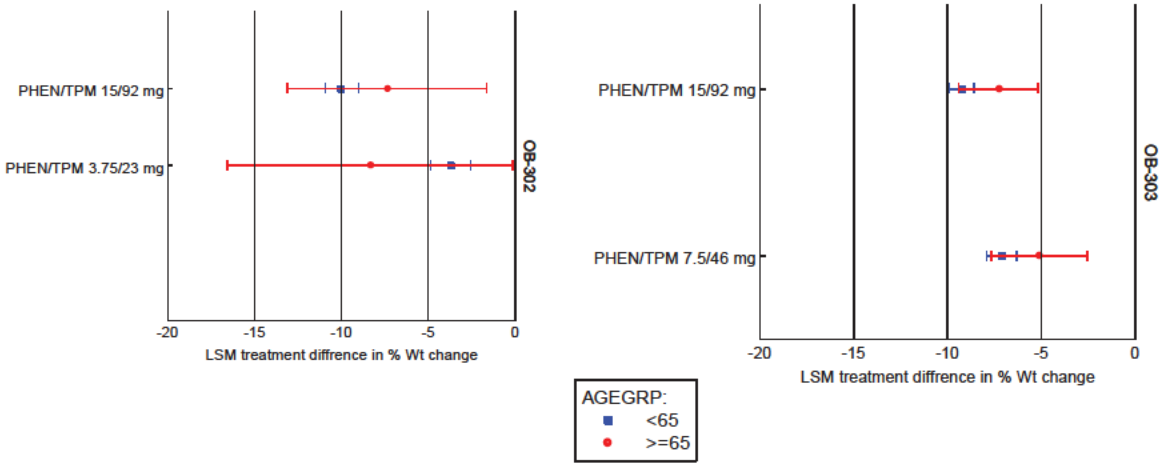


Race

Treatment-by-race interaction was not significant (p=0.94).

Age group

In Study 303, treatment-by-age group (<65) was significant for the PHEN/TPM 15/92 mg group (p=0.08) and the PHEN/TPM 7.5/46 mg group (p=0.09). Patients <65 years of age lost a mean of 2% more body weight than patients aged ≥65. The interaction was not significant for Study 302 (p=0.4) (PHEN/TPM 15/92 mg). The interaction was quantitative in nature.



5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

For efficacy, PHEN/TPM 15/92 mg and 7.5/46 mg were superior to both the respective components and placebo in the 26 week factorial study in weight reduction. The Efficacy of PHEN/TPM 3.75/23 mg was not established over and above the respective components since the dose was not included in the factorial study.

5.2 Conclusions and Recommendations

Based on the analysis of efficacy and selected safety data, we recommend that 7.5/46 mg Qnexa should be the indicated dose.

APPENDICES

1 Quality of Life

The IWQOL-Lite is a 31-item, self-administered instrument to evaluate physical function, self-esteem, sexual life, public distress, and work. The IWQOL scores were mapped to a 0 to 100 scale by subtracting the observed score from the maximum value for the component, dividing by the range for the component, and multiplying by 100.

Patients completed the IWQOL-Lite questionnaire at screening and study completion (Visit 10, week 28 or early termination).

One of the following secondary efficacy variables in Study 301 was patient reported outcome (PRO):

- Percentage of patients with at least 10% weight loss at Week 28 with LOCF,
- Change in waist circumference from baseline to Week 28 with LOCF, (key 2nd) and
- Changes in impact of Weight on Quality of Life (IWQOL)-Lite composite and individual domain scores at Week 28 with LOCF.

The PRO was an exploratory variable for studies 302 and 303.

Analysis of IWQOL composite and domain scores used the same ANCOVA model for percent weight loss at week 28 with LOCF. The IWQOL scores were mapped to a 0 to 100 scale by subtracting the observed score from the maximum value for the component, dividing by the range for the component, and multiplying by 100.

According to the sponsor 'The IWQOL data were re-scaled to handle missing data and to map them onto the more commonly used 0 to 100 scale rather than the minimum to maximum raw score scaling that was originally analyzed. Also, the original analysis required all components to be answered to be considered for analysis, whereas the re-scaled scores were computed if a minimum of 50% of the items for each component were answered and a minimum of 75% of all items were answered to compute a valid total score. To handle missing data, the mean of the observed components was first calculated and then this was multiplied by the number of items in each component to provide the expected score that would have been observed had all items been answered. To map to the 0 to 100 scale, the re-scaled score was subtracted from the maximum value for that component, divided by the range of the component, and then multiplied by 100. Re-scaling was performed to be consistent with the industry standard for interpretation and analysis of IWQOL data and the analysis tables that were described in the original SAP were re-generated using the re-scaled scores as input.'

The analyses of IWQOL however, included patients providing complete responses to each domain or composite total; i.e., no imputation will be performed to estimate missing data for this variable.

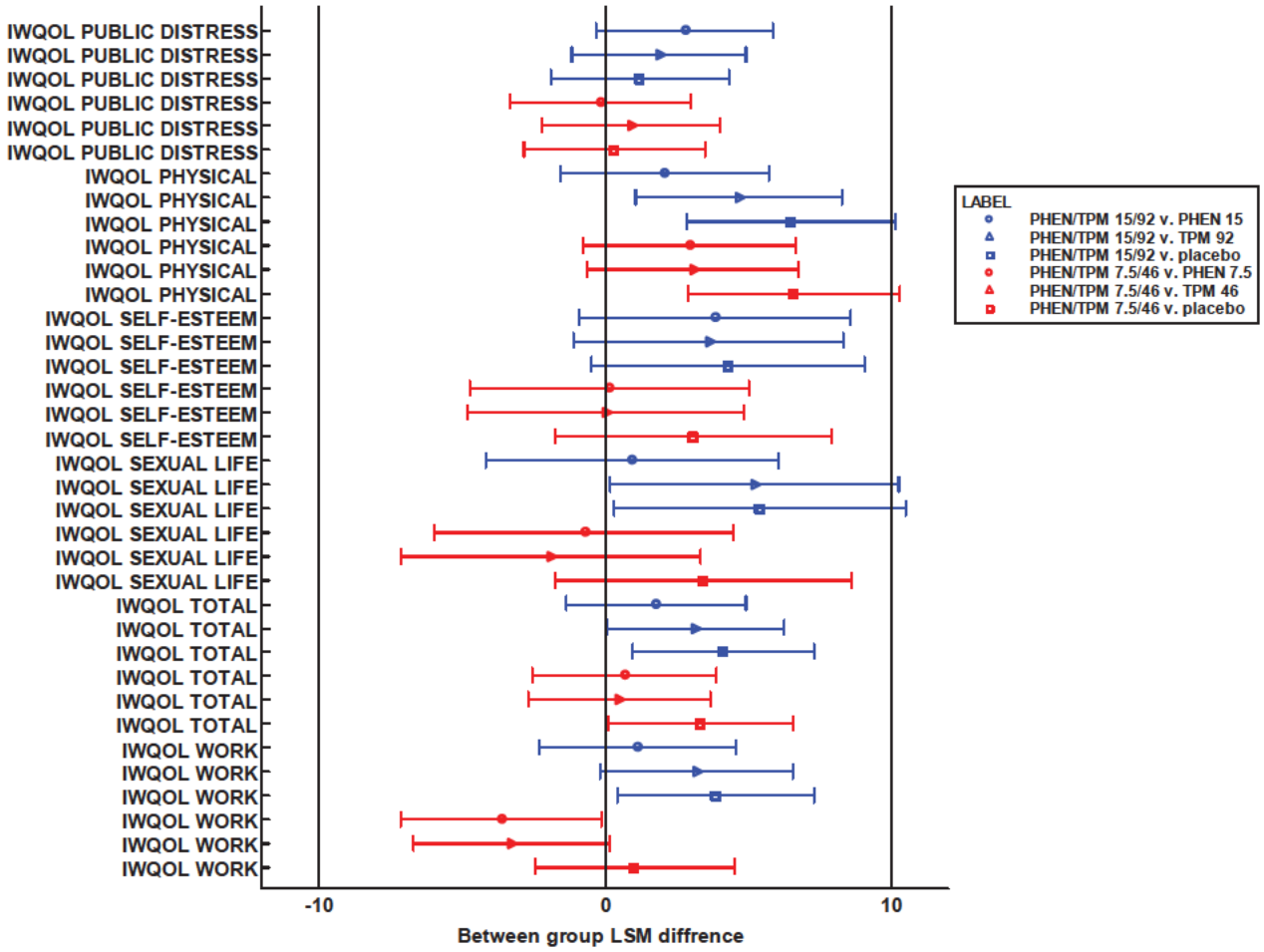
Change from baseline to week 28 or LOCF used an ANCOVA model with treatment and gender as fixed effects and baseline as covariate.

Table 1 displays the statistical analysis results for the IWQOL. For all IWQOL domains, none of the combinations were statistically significant to both of the components. For self-esteem, sexual life, public distress and work none of the comparisons (vs. placebo or components) were statistically significant. There were a total of 36 comparisons, the multiplicity need to be addressed in advance. The PRO guidance stated that 'Sponsors should avoid cherry picking or post hoc selective picking of PRO endpoint results for inclusion in proposed labeling.'

Table 1 ANCOVA results for IWQOL by treatment

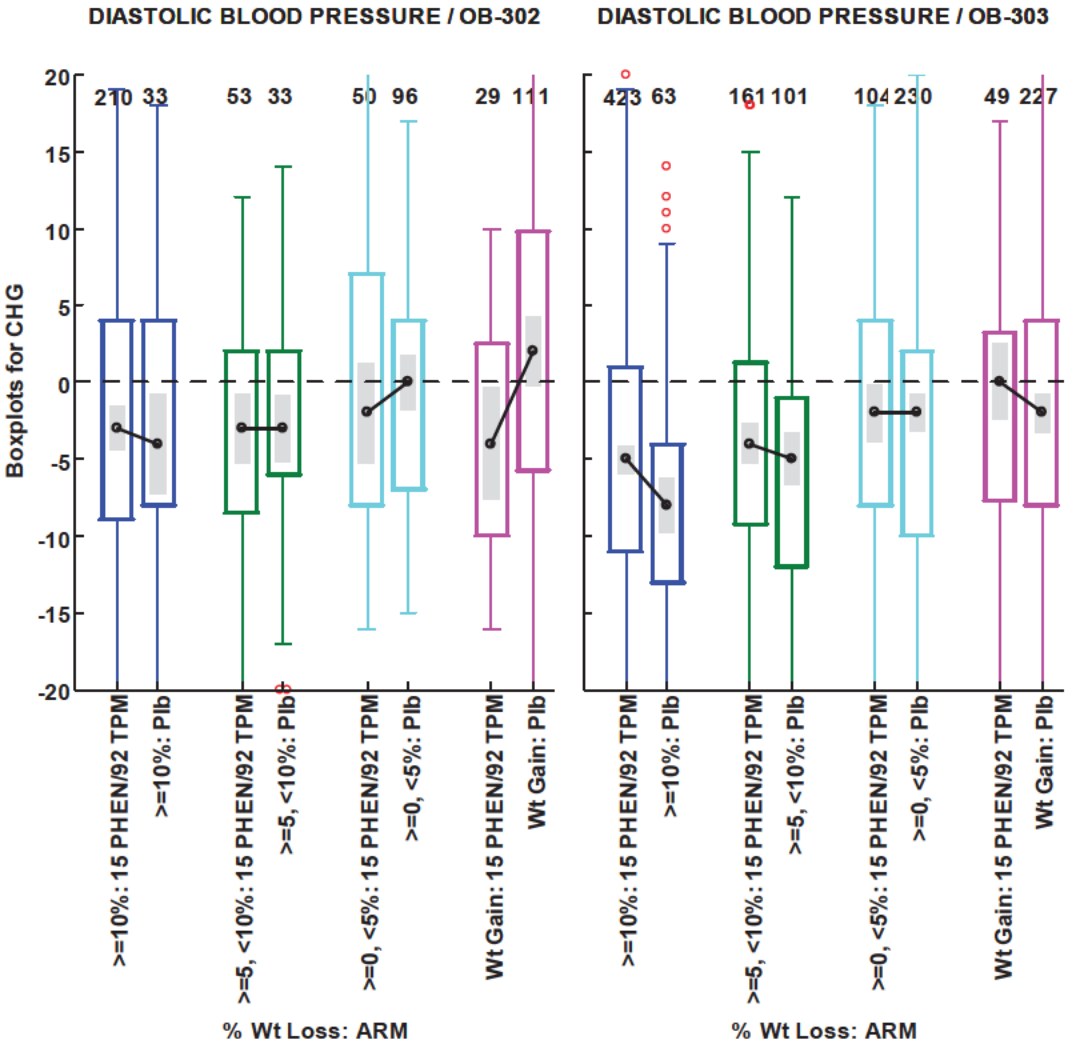
		Treatment						
IWQOL		15P/92 T n=96	15P n=87	92T n=91	7.5P/46T n=87	7.5P n=83	46T n=87	Plb n=85
Physical function	LSM	12.6	11.4	7.3	13.0	11.4	9.0	6.2
	SE	(1.6)	(1.6)	(1.6)	(1.6)	(1.6)	(1.6)	(1.6)
	p-value	vs. plb	comb. vs. comp		vs. plb	comb. vs. comp		
		<0.01	**0.60**	**0.01**	**<0.01**	**0.48**	**0.08**	
Self-esteem	LSM	15.1	12.3	10.4	11.6	13.8	11.4	9.8
	SE	(2.0)	(2.1)	(2.1)	(2.1)	(2.1)	(2.1)	(2.1)
	p-value	vs. plb	vs. comp		vs. plb	vs. comp		
		0.06	**0.33**	**0.10**	**0.53**	**0.47**	**0.94**	
Sexual life	LSM	12.2	11.2	6.6	8.7	10.6	9.7	6.4
	SE	(2.1)	(2.3)	(2.2)	(2.3)	(2.2)	(2.3)	(2.2)
	p-value	vs. plb	vs. component		vs. plb	vs. component		
		0.06	**0.75**	**0.07**	**0.47**	**0.55**	**0.76**	
Public distress	LSM	6.8	4.5	4.7	4.8	5.8	4.7	4.8
	SE	(1.3)	(1.4)	(1.4)	(1.4)	(1.4)	(1.4)	(1.4)
	p-value	vs. plb	vs. component	vs. plb	vs. comp.			
		0.30	**0.23**	**0.27**	**1.0**	**0.6**	**1.0**	
Work	LSM	7.6	5.7	3.6	4.2	7.6	7.0	3.7
	SE	(1.4)	(1.5)	(1.5)	(1.5)	(1.5)	(1.5)	(1.5)
	p-value	vs. plb	vs. component	vs. plb	vs. comp.			
		0.06	**0.37**	**0.055**	**0.79**	**0.12**	**0.19**	
Composite (Total)	LSM	11.2	9.9	7.2	9.4	10.2	8.6	6.5
	SE	(1.3)	(1.4)	(1.4)	(1.4)	(1.4)	(1.4)	(1.4)
	p-value	vs. plb	vs. component	vs. plb	vs. comp.			
		0.01	**0.48**	**0.04**	**0.13**	**0.68**	**0.66**	

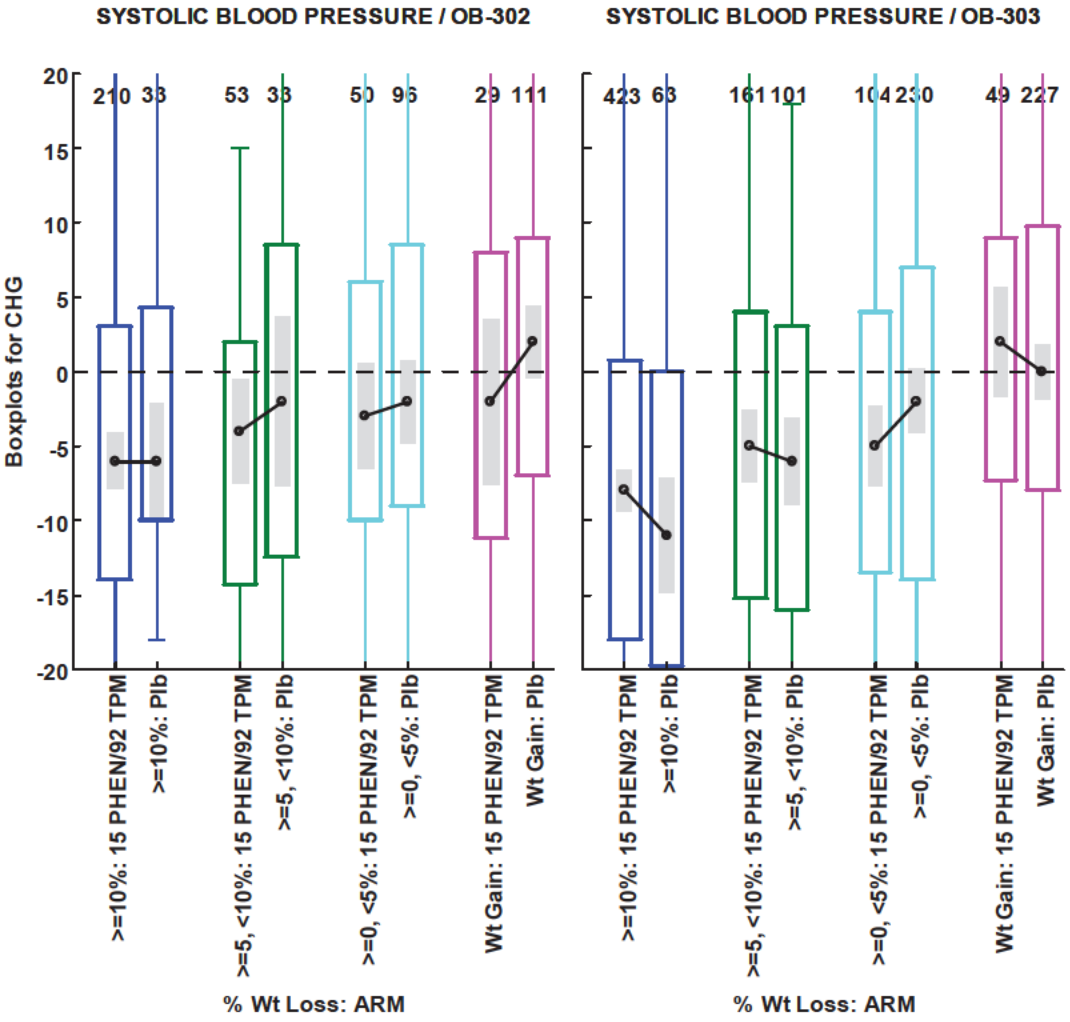
Figure 20 Between-group least squared mean difference (95% CI) – Study 301

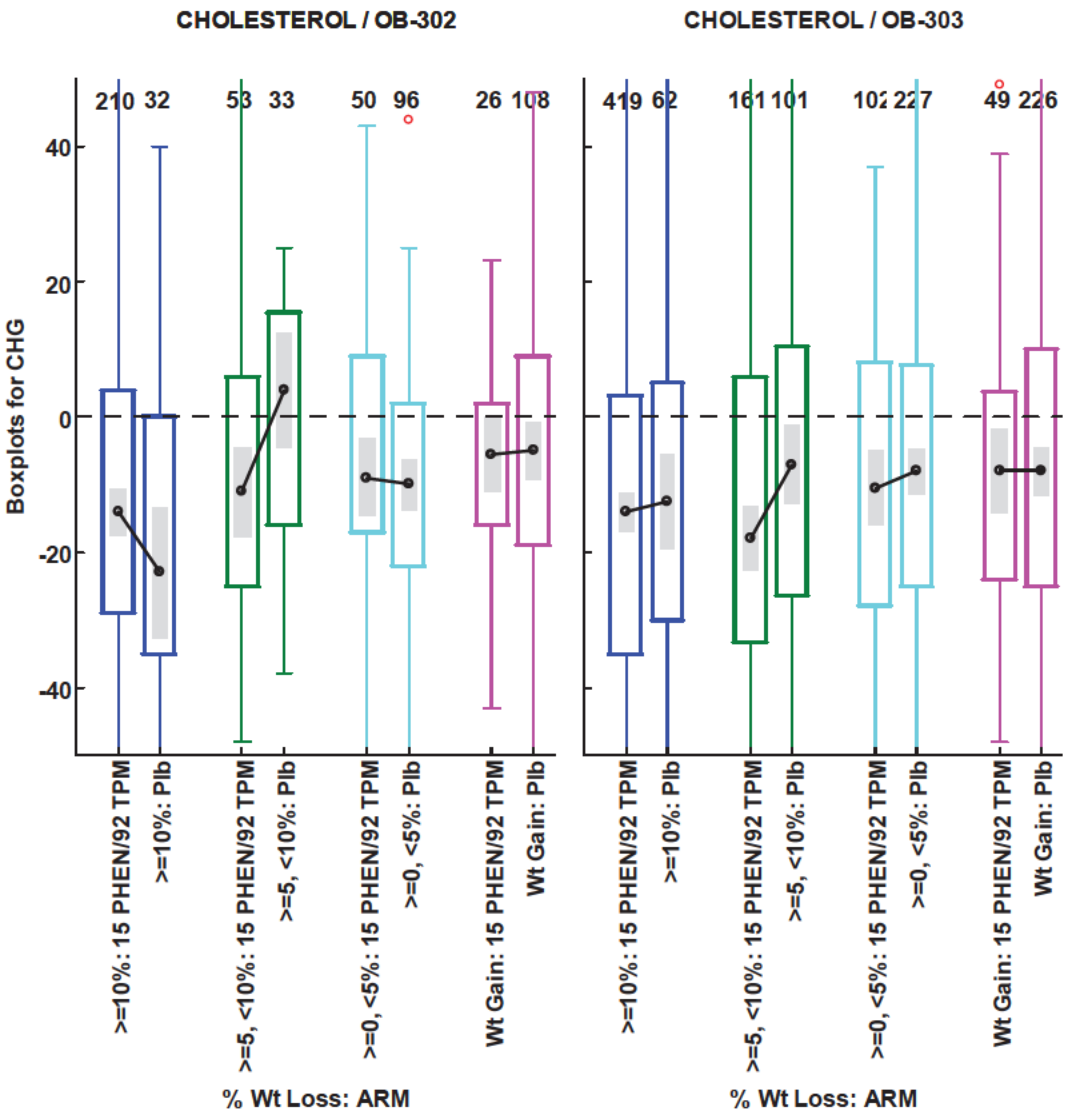


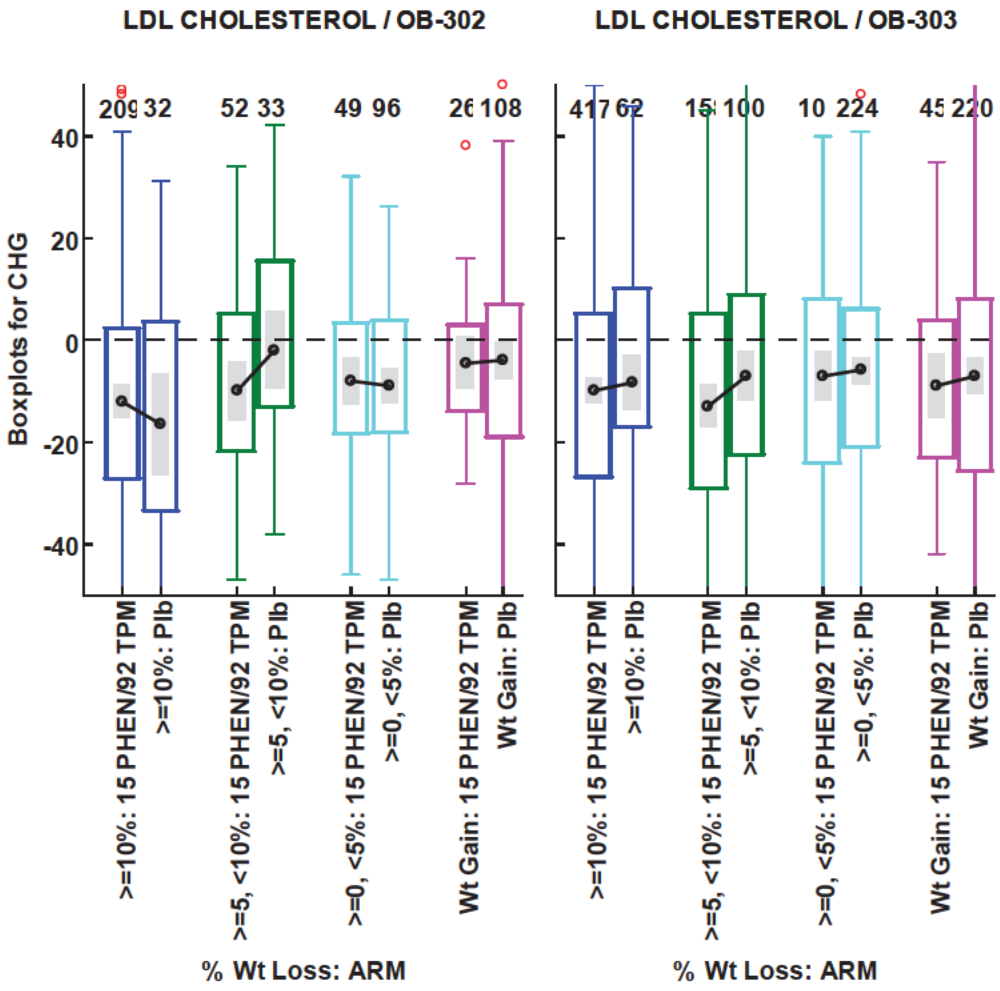
2. Boxplots by weight loss category and treatment group

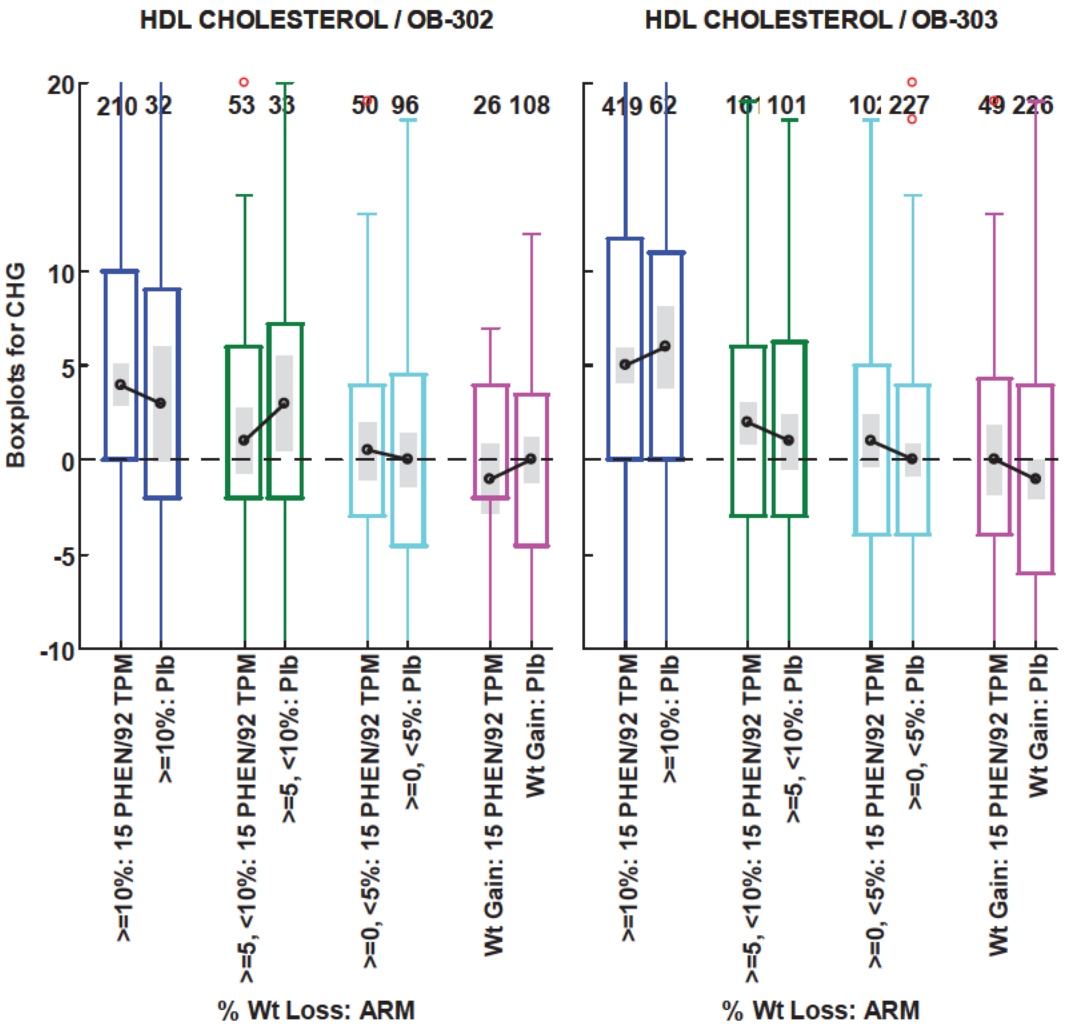
As requested by the Medical Division, the following are boxplots for change from baseline to week 56 (on drug MITT Completers) in co-morbidities by treatment groups (15 mg PHEN/92 mg TPM and placebo) and percent weight loss categories (≥10%, ≥5% and <10%, ≥0% and <5%, and weight gain) at week 56 (on drug MITT Completers)/

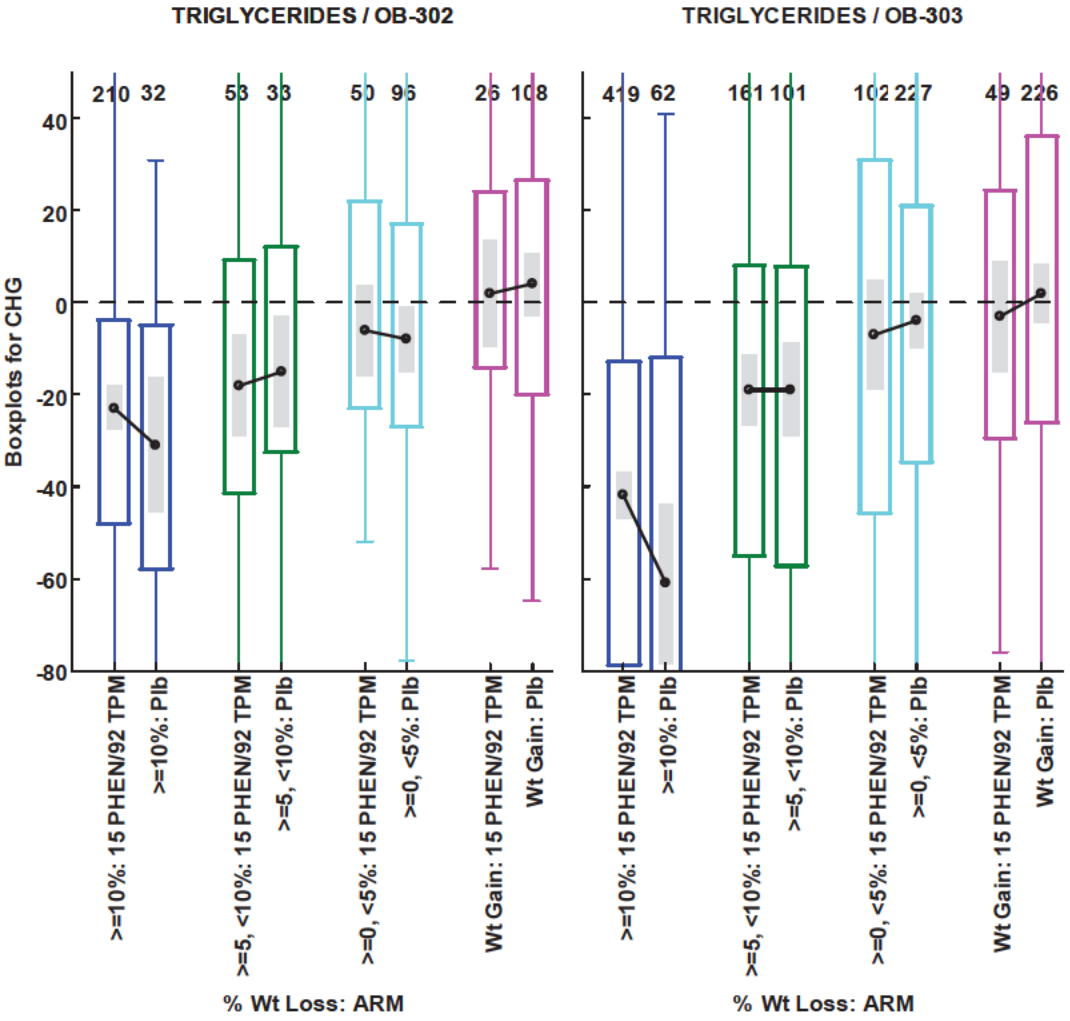


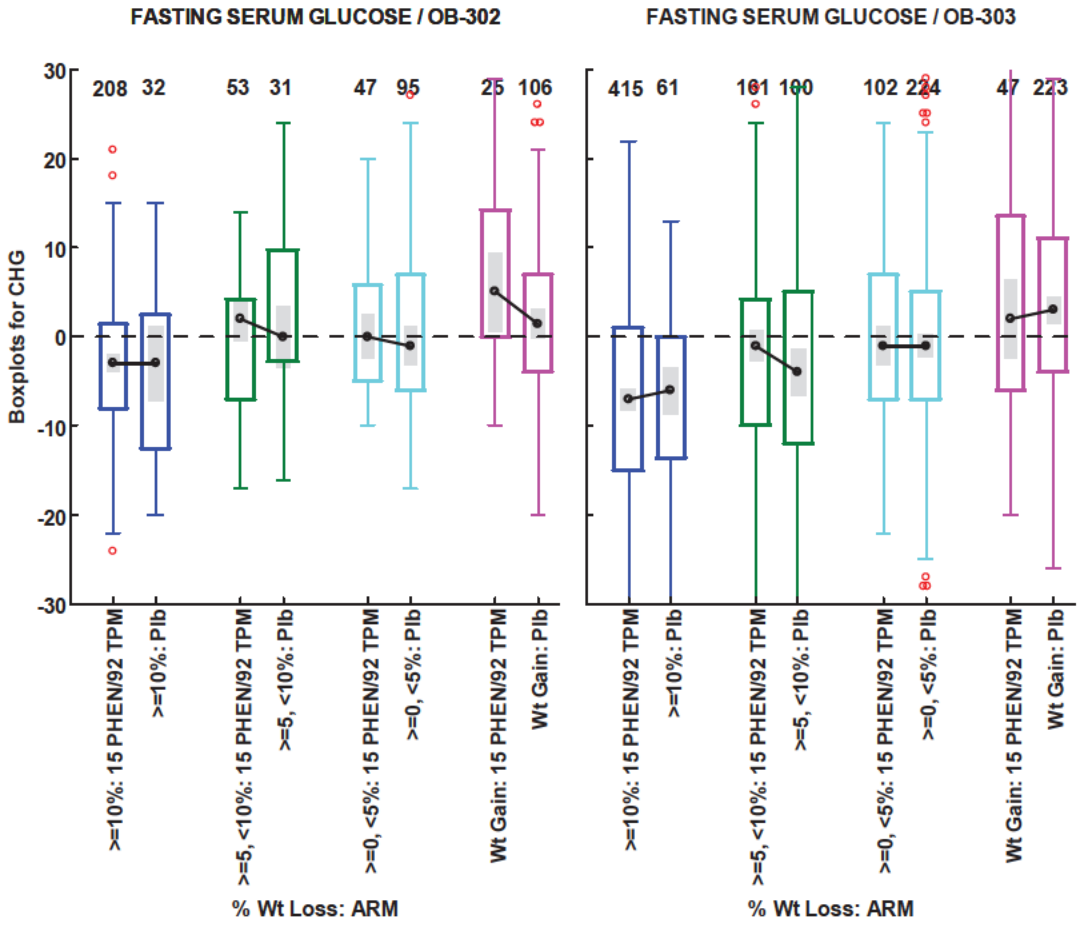

SYSTOLIC BLOOD PRESSURE / OB-302
SYSTOLIC BLOOD PRESSURE / OB-303
Boxplots for CHG
20
15
10
5
0
-5
-10
-15
-20
210 33
53 33
50 96
29 111
423 63
161 101
104 230
49 227
>=10%: 15 PHEN/92 TPM
>=10%: Plb
>=5, <10%: 15 PHEN/92 TPM
>=5, <10%: Plb
>=0, <5%: 15 PHEN/92 TPM
>=0, <5%: Plb
Wt Gain: 15 PHEN/92 TPM
Wt Gain: Plb
% Wt Loss: ARM
% Wt Loss: ARM

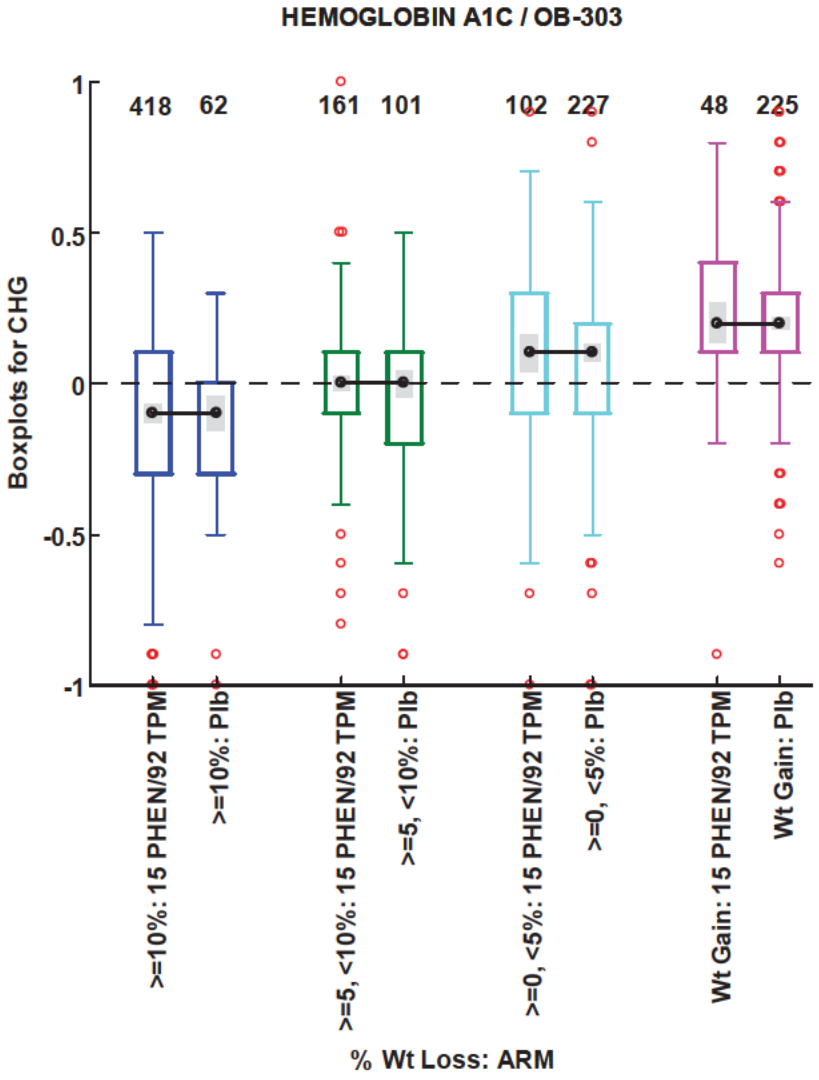

CHOLESTEROL / OB-302
CHOLESTEROL / OB-303
Boxplots for CHG
40
20
0
-20
-40
210 32
53 33
50 96
26 108
419 62
161 101
102 227
49 226
>=10%: 15 PHEN/92 TPM
>=10%: Plb
>=5, <10%: 15 PHEN/92 TPM
>=5, <10%: Plb
>=0, <5%: 15 PHEN/92 TPM
>=0, <5%: Plb
Wt Gain: 15 PHEN/92 TPM
Wt Gain: Plb
% Wt Loss: ARM
% Wt Loss: ARM

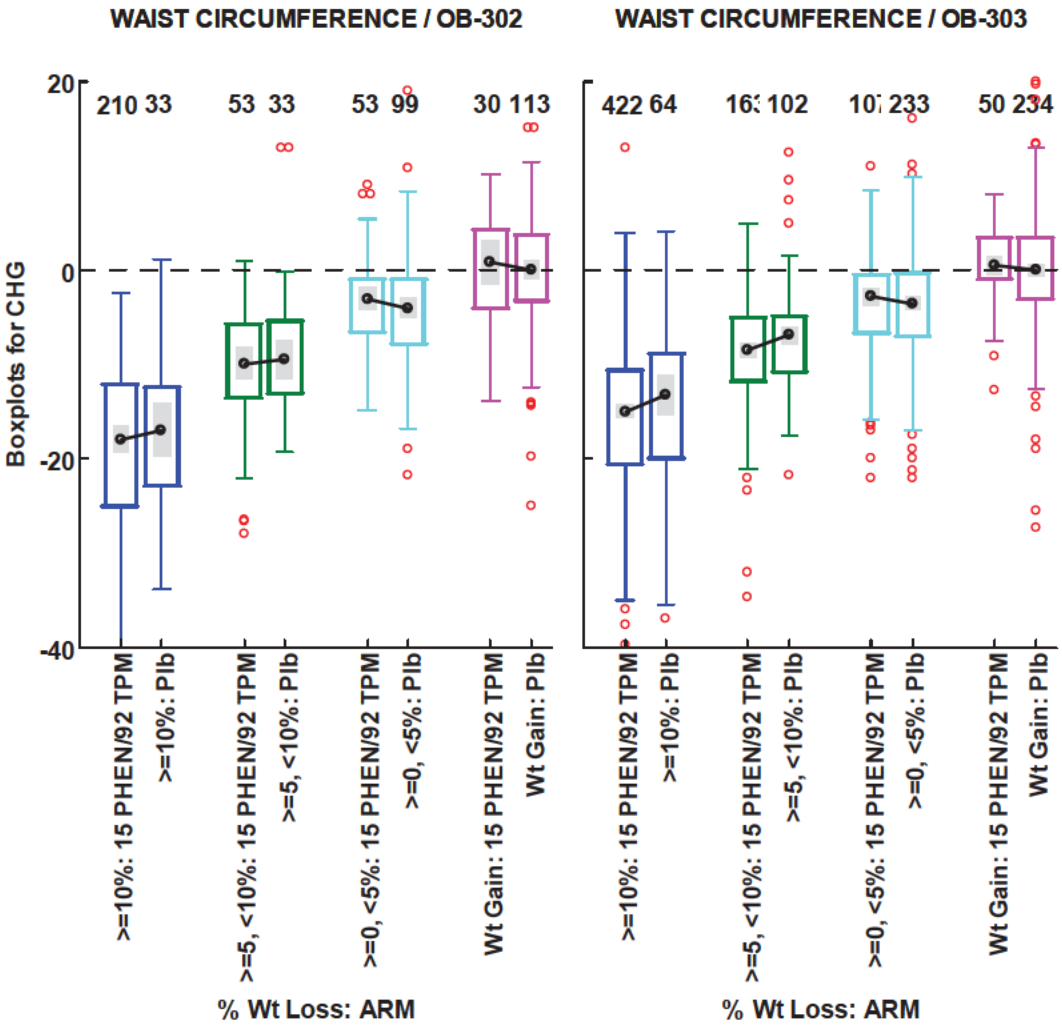

LDL CHOLESTEROL / OB-302
LDL CHOLESTEROL / OB-303
Boxplots for CHG
40
20
0
-20
-40
209 32
52 33
49 96
26 108
417 62
15 100
10 224
45 220
>=10%: 15 PHEN/92 TPM
>=10%: Plb
>=5, <10%: 15 PHEN/92 TPM
>=5, <10%: Plb
>=0, <5%: 15 PHEN/92 TPM
>=0, <5%: Plb
Wt Gain: 15 PHEN/92 TPM
Wt Gain: Plb
% Wt Loss: ARM
% Wt Loss: ARM

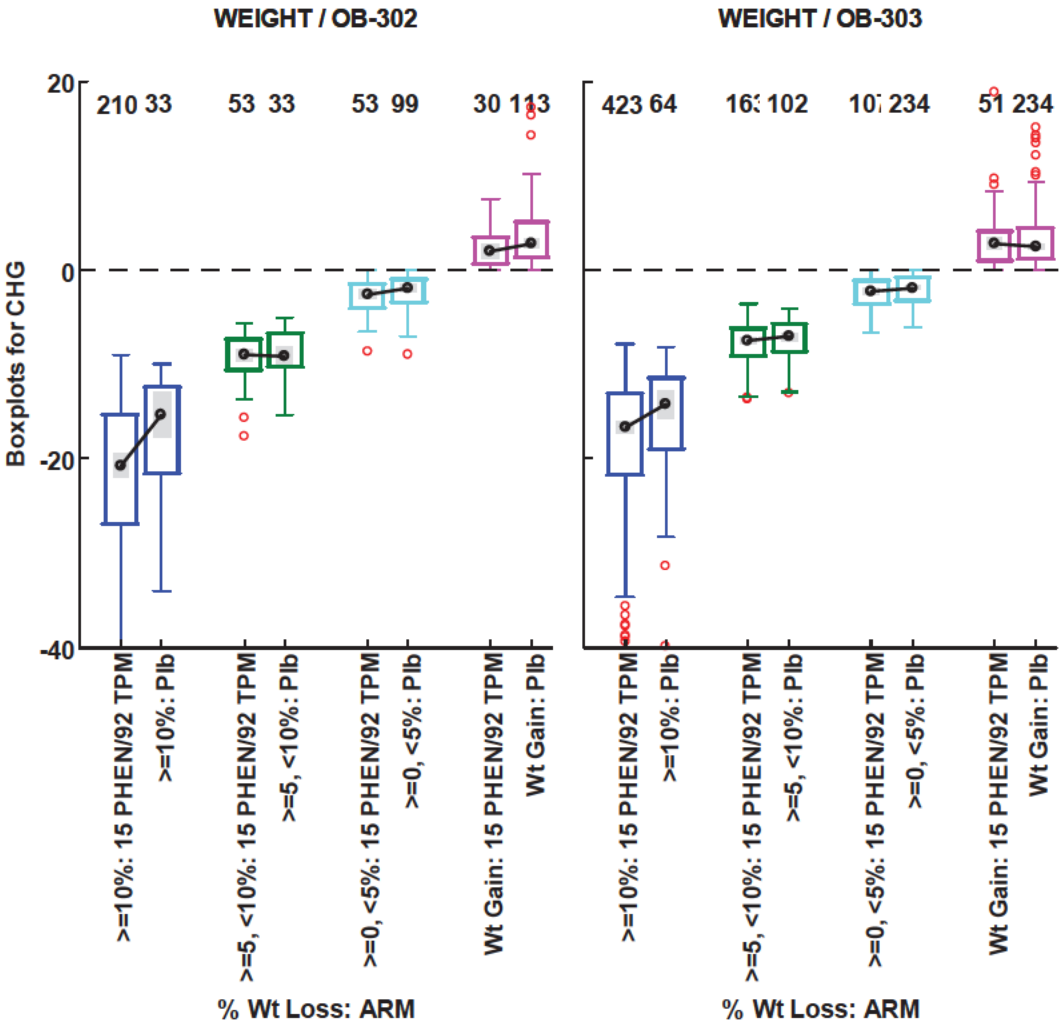

HDL CHOLESTEROL / OB-302
HDL CHOLESTEROL / OB-303
Boxplots for CHG
20
10
5
0
-5
-10
210 32
53 33
50 96
26 108
419 62
101
227
49
>=10%: 15 PHEN/92 TPM
>=10%: Plb
>=5, <10%: 15 PHEN/92 TPM
>=5, <10%: Plb
>=0, <5%: 15 PHEN/92 TPM
>=0, <5%: Plb
Wt Gain: 15 PHEN/92 TPM
Wt Gain: Plb
% Wt Loss: ARM
% Wt Loss: ARM

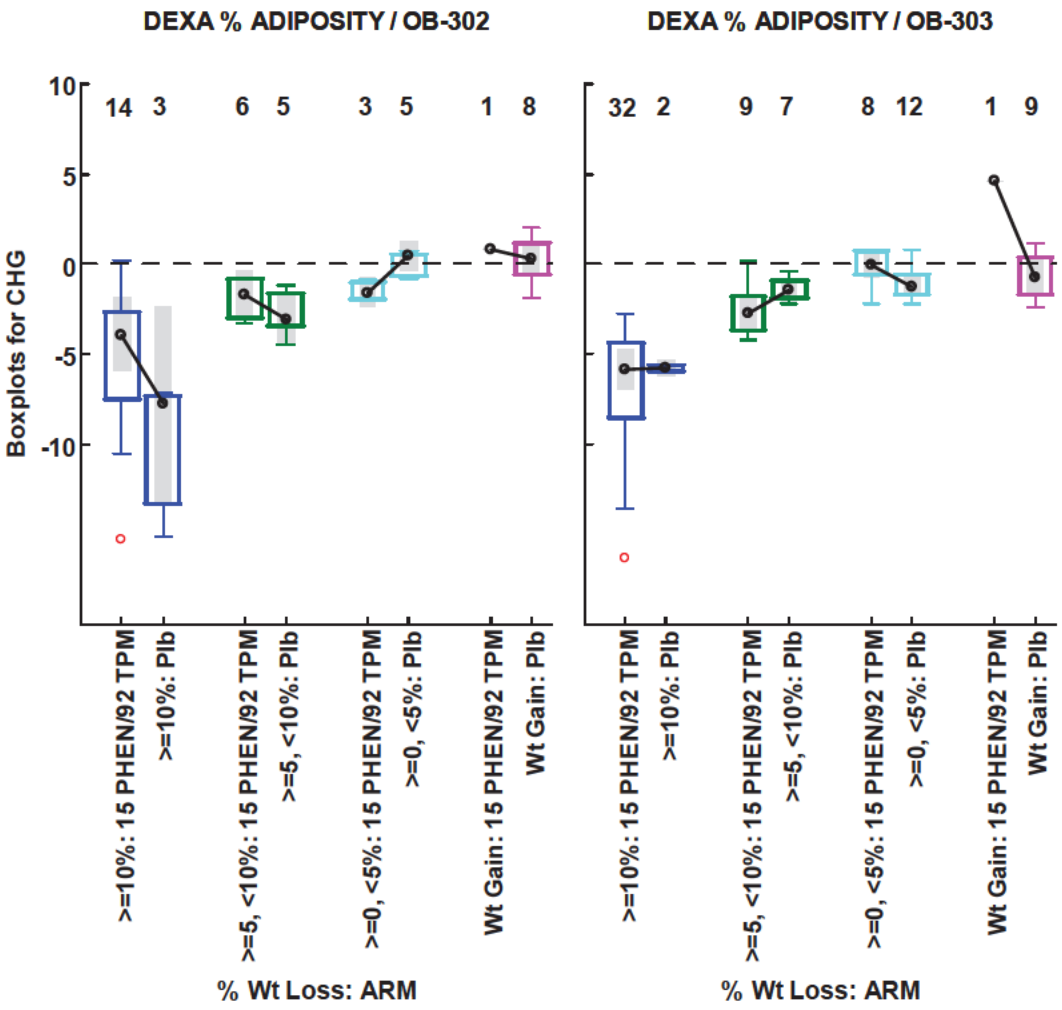

TRIGLYCERIDES / OB-302
TRIGLYCERIDES / OB-303
Boxplots for CHG
40
20
0
-20
-40
-60
-80
210 32
53 33
50 96
26 108
419 62
161 101
102 227
49 226
>=10%: 15 PHEN/92 TPM
>=10%: Plb
>=5, <10%: 15 PHEN/92 TPM
>=5, <10%: Plb
>=0, <5%: 15 PHEN/92 TPM
>=0, <5%: Plb
Wt Gain: 15 PHEN/92 TPM
Wt Gain: Plb
% Wt Loss: ARM
% Wt Loss: ARM

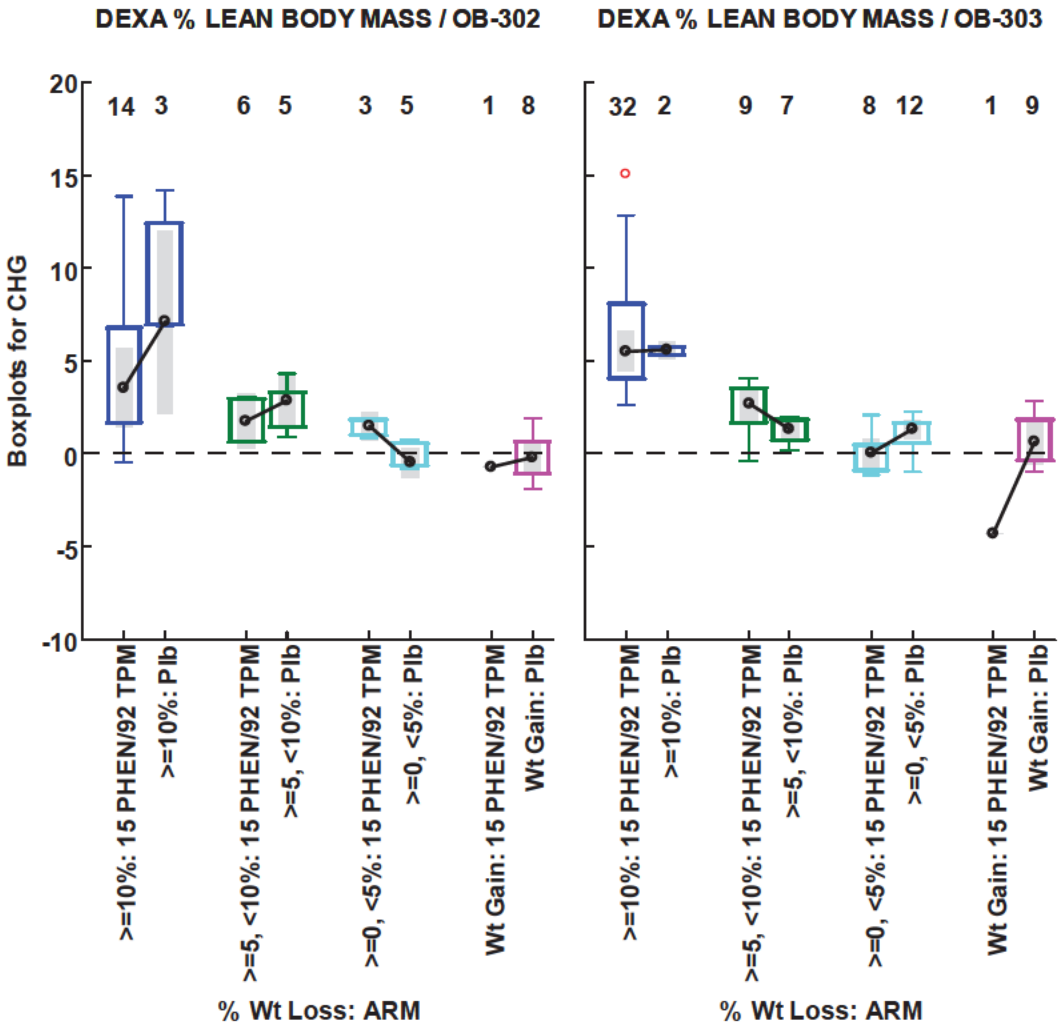

FASTING SERUM GLUCOSE / OB-302
FASTING SERUM GLUCOSE / OB-303
Boxplots for CHG
30
20
10
0
-10
-20
-30
208 32 53 31 47 95 25 106
415 61 161 160 102 224 47 223
>=10%: 15 PHEN/92 TPM
>=10%: Plb
>=5, <10%: 15 PHEN/92 TPM
>=5, <10%: Plb
>=0, <5%: 15 PHEN/92 TPM
>=0, <5%: Plb
Wt Gain: 15 PHEN/92 TPM
Wt Gain: Plb
% Wt Loss: ARM


HEMOGLOBIN A1C / OB-303
Boxplots for CHG
1
0.5
0
-0.5
-1
418
62
161
101
102
227
48
225
>=10%: 15 PHEN/92 TPM
>=10%: Plb
>=5, <10%: 15 PHEN/92 TPM
>=5, <10%: Plb
>=0, <5%: 15 PHEN/92 TPM
>=0, <5%: Plb
Wt Gain: 15 PHEN/92 TPM
Wt Gain: Plb
% Wt Loss: ARM


WAIST CIRCUMFERENCE / OB-302
WAIST CIRCUMFERENCE / OB-303
Boxplots for CHG
20
0
-20
-40
210 33
53 33
53 99
30 113
422 64
163 102
107 233
50 234
>=10%: 15 PHEN/92 TPM
>=10%: Plb
>=5, <10%: 15 PHEN/92 TPM
>=5, <10%: Plb
>=0, <5%: 15 PHEN/92 TPM
>=0, <5%: Plb
Wt Gain: 15 PHEN/92 TPM
Wt Gain: Plb
% Wt Loss: ARM
% Wt Loss: ARM

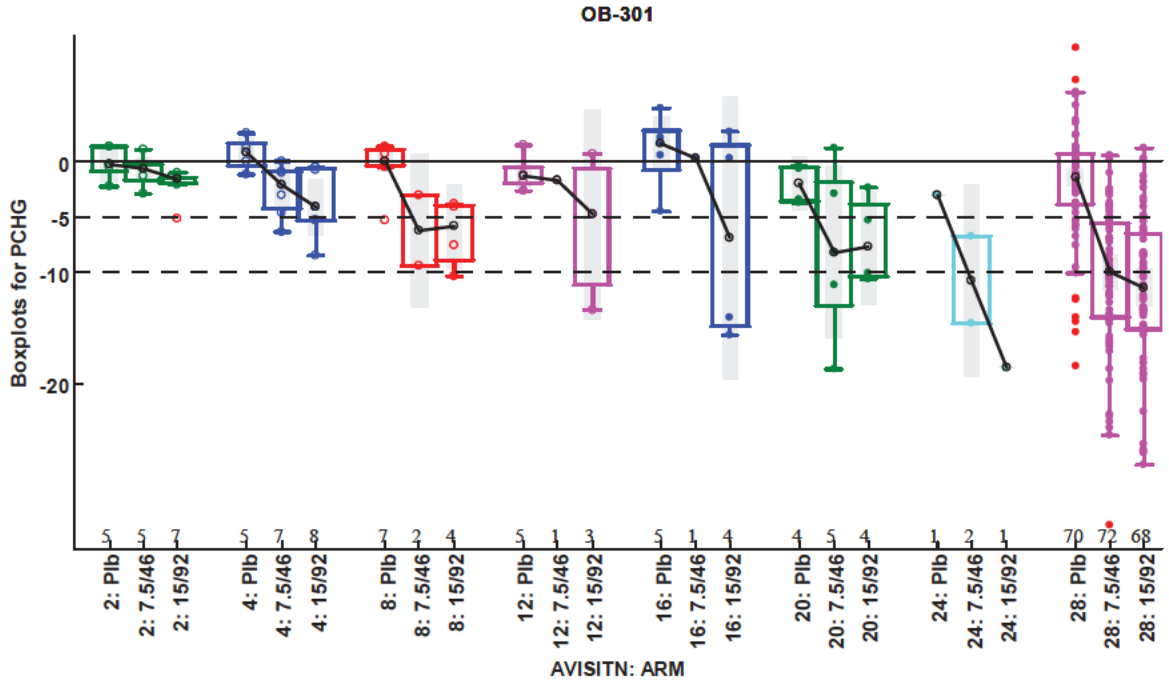

WEIGHT / OB-302
WEIGHT / OB-303
Boxplots for CHG
20
0
-20
-40
210 33
53 33
53 99
30 113
423 64
163 102
107 234
51 234
>=10%: 15 PHEN/92 TPM
>=10%: Plb
>=5, <10%: 15 PHEN/92 TPM
>=5, <10%: Plb
>=0, <5%: 15 PHEN/92 TPM
>=0, <5%: Plb
Wt Gain: 15 PHEN/92 TPM
Wt Gain: Plb
% Wt Loss: ARM

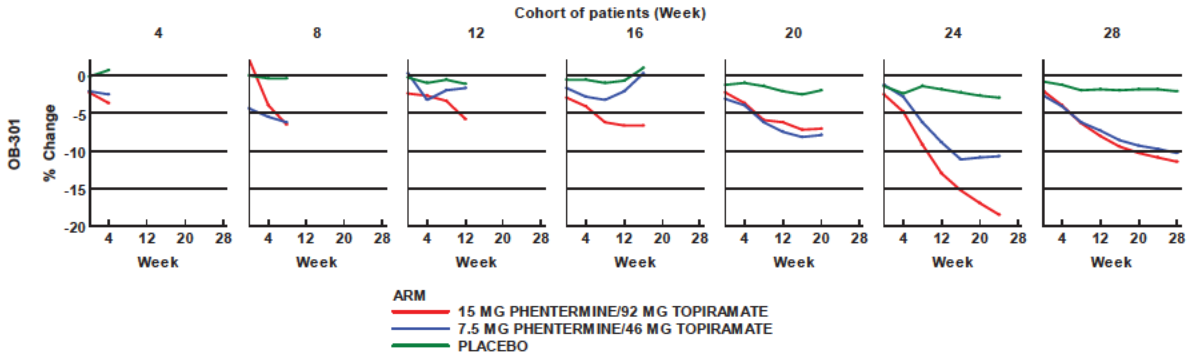

DEXA % ADIPOSITY / OB-302
DEXA % ADIPOSITY / OB-303
Boxplots for CHG
10
5
0
-5
-10
14 3 6 5 3 5 1 8
32 2 9 7 8 12 1 9
>=10%: 15 PHEN/92 TPM
>=10%: Plb
>=5, <10%: 15 PHEN/92 TPM
>=5, <10%: Plb
>=0, <5%: 15 PHEN/92 TPM
>=0, <5%: Plb
Wt Gain: 15 PHEN/92 TPM
Wt Gain: Plb
% Wt Loss: ARM
% Wt Loss: ARM

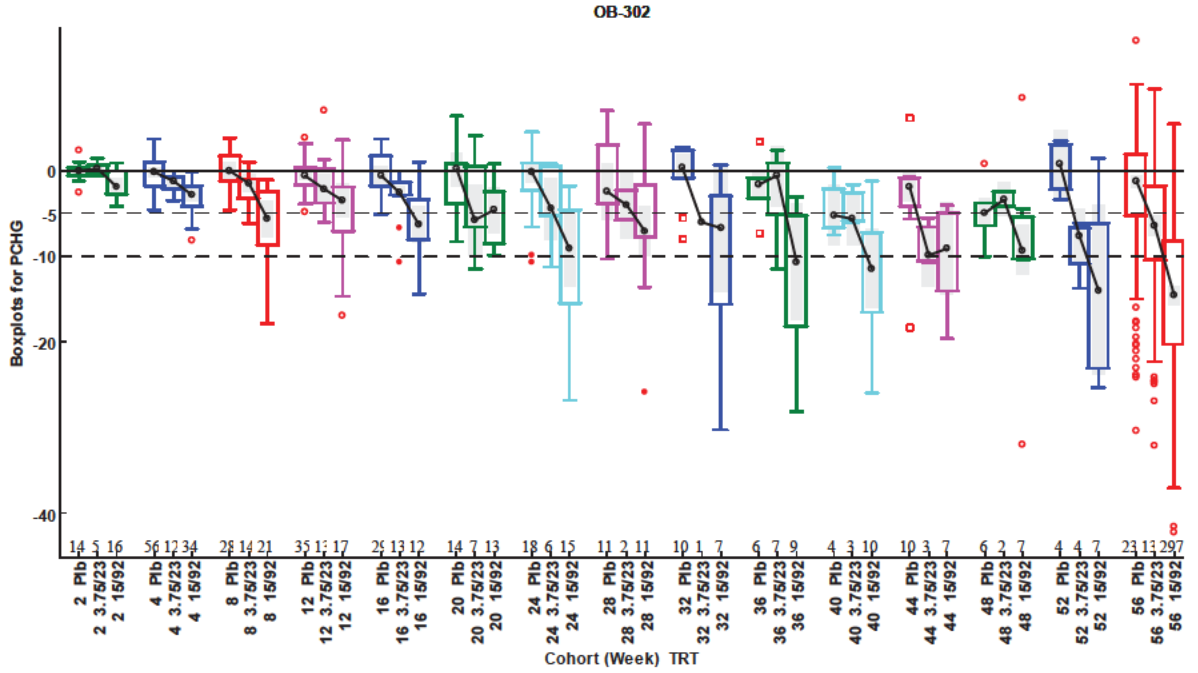

DEXA % LEAN BODY MASS / OB-302
DEXA % LEAN BODY MASS / OB-303
Boxplots for CHG
20
15
10
5
0
-5
-10
14 3 6 5 3 5 1 8
32 2 9 7 8 12 1 9
>=10%: 15 PHEN/92 TPM
>=10%: Plb
>=5, <10%: 15 PHEN/92 TPM
>=5, <10%: Plb
>=0, <5%: 15 PHEN/92 TPM
>=0, <5%: Plb
Wt Gain: 15 PHEN/92 TPM
Wt Gain: Plb
% Wt Loss: ARM
% Wt Loss: ARM

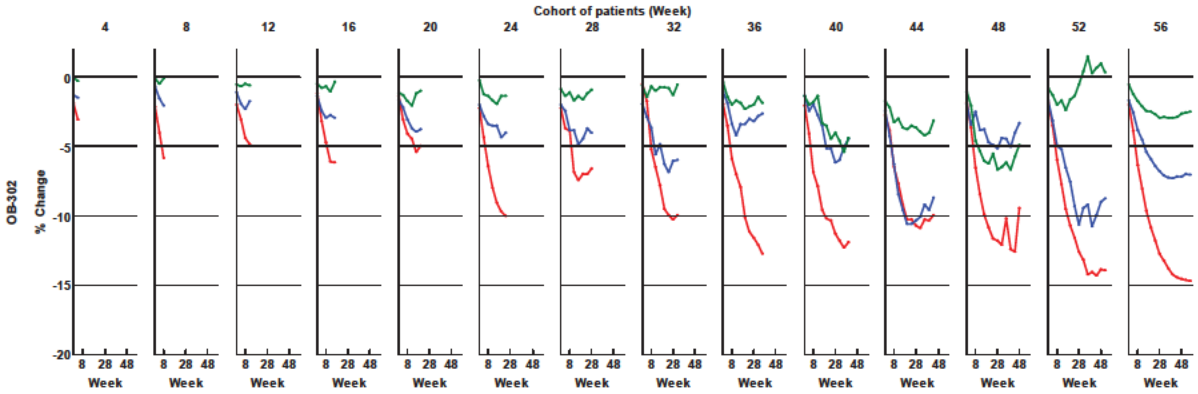
3. Weight change in discontinuation cohort of patients

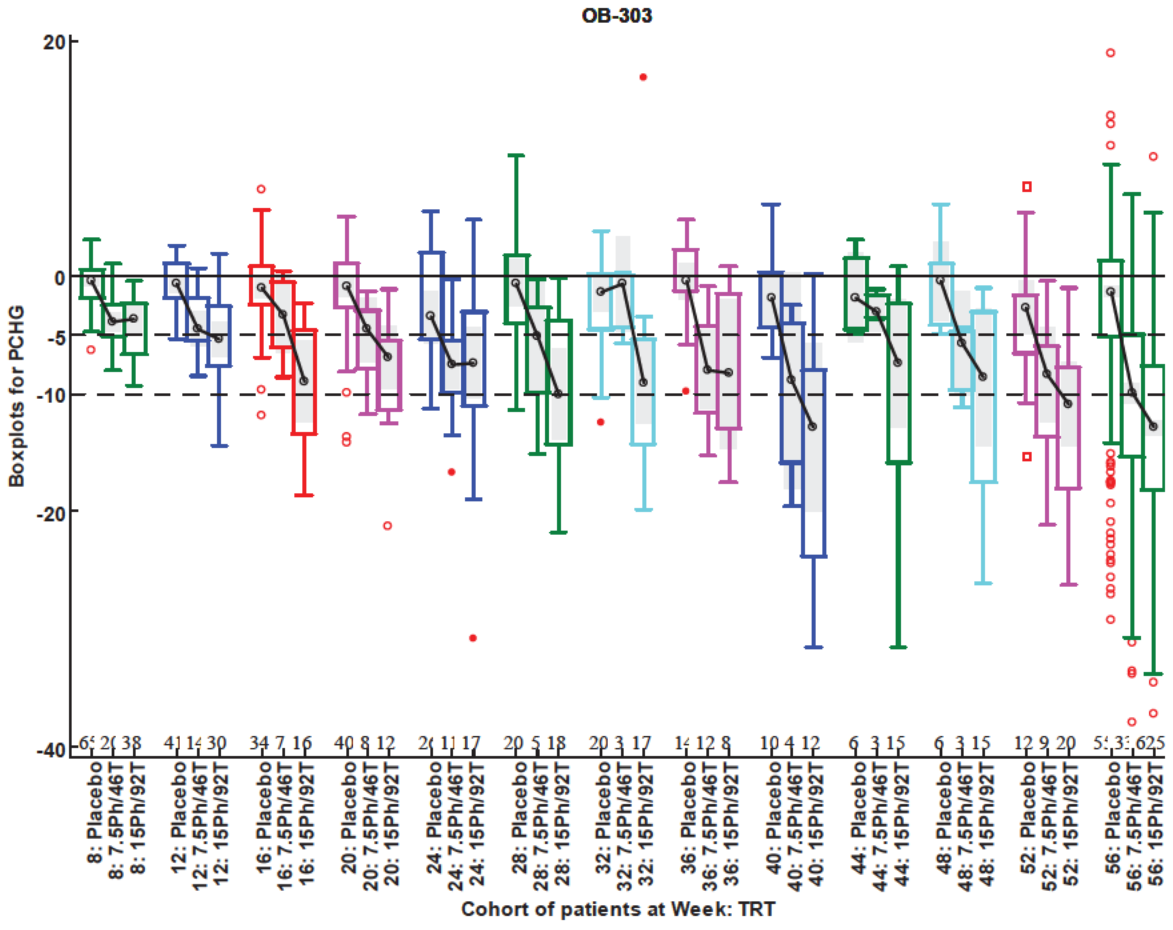
Approximately 40% of patients discontinued study medication. The following graphs show different cohorts of patients defined by their last week on study. The graphs showed that the cohorts who discontinued medication were similar to the completers (last cohort) in trend of percent weight change from baseline.

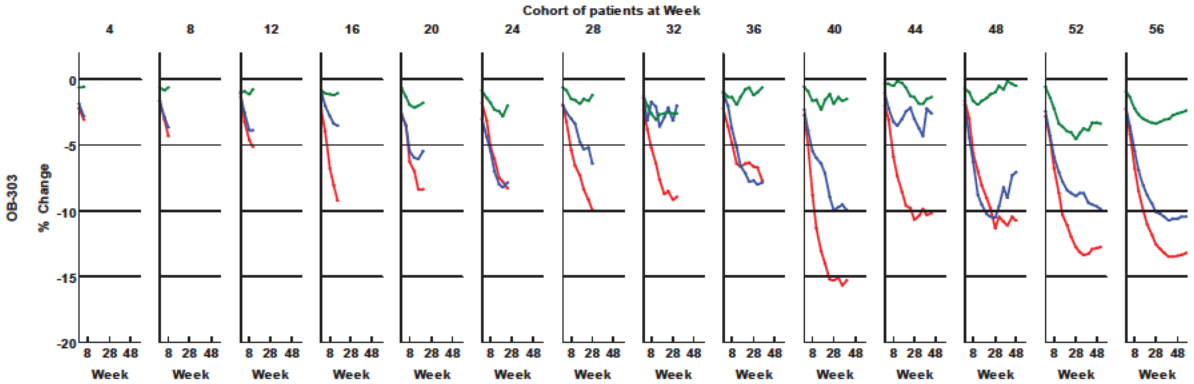





OB-302
Boxplots for PCHG
0
-5
-10
-20
-40
14 5 16 56 12 34 28 14 21 35 15 17 25 13 12 14 7 13 18 6 15 11 2 11 10 1 7 6 7 9 4 3 10 10 3 7 6 2 7 4 4 7 23 13 29 7 8
2 Plb
2 3.75/23
2 15/92
4 Plb
4 3.75/23
4 15/92
8 Plb
8 3.75/23
8 15/92
12 Plb
12 3.75/23
12 15/92
16 Plb
16 3.75/23
16 15/92
20 Plb
20 3.75/23
20 15/92
24 Plb
24 3.75/23
24 15/92
28 Plb
28 3.75/23
28 15/92
32 Plb
32 3.75/23
32 15/92
36 Plb
36 3.75/23
36 15/92
40 Plb
40 3.75/23
40 15/92
44 Plb
44 3.75/23
44 15/92
48 Plb
48 3.75/23
48 15/92
52 Plb
52 3.75/23
52 15/92
56 Plb
56 3.75/23
56 15/92
Cohort (Week) TRT

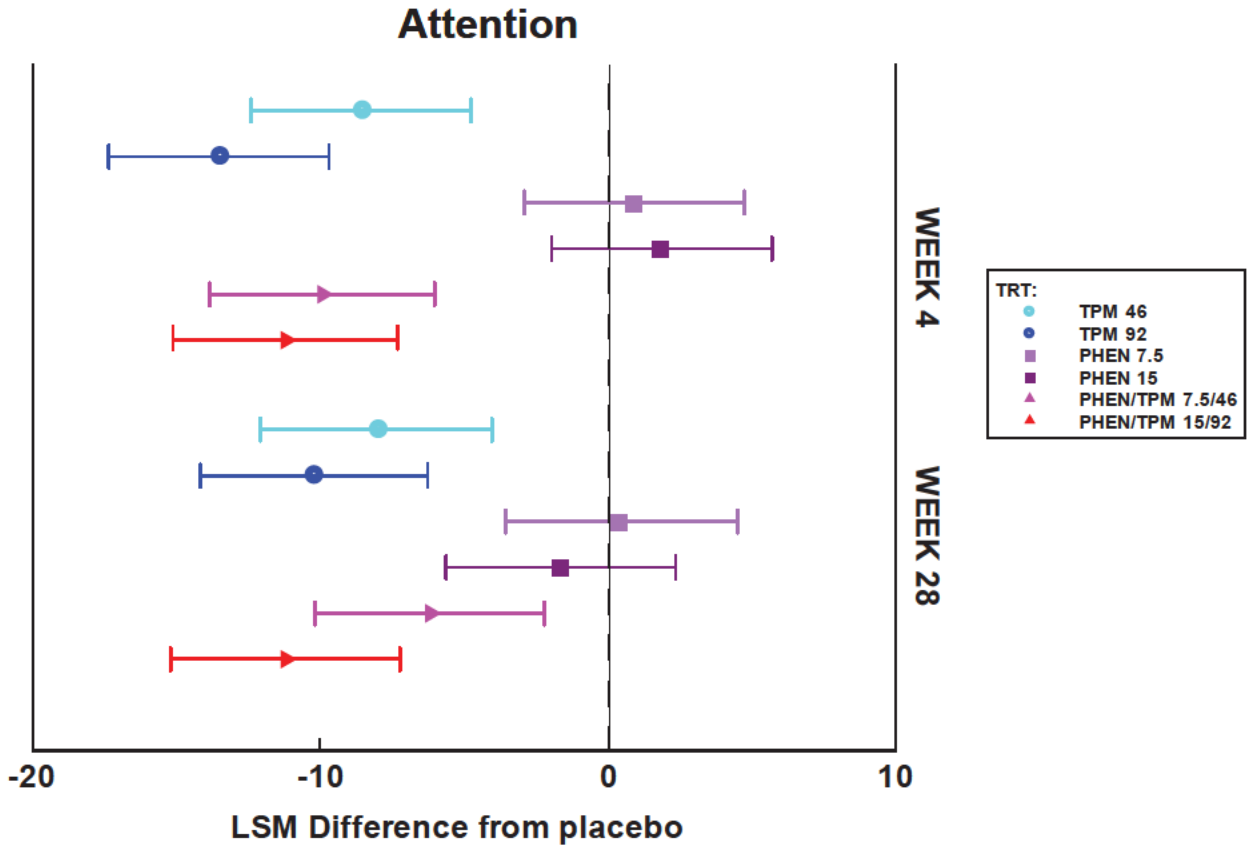

Cohort of patients (Week)
4
8
12
16
20
24
28
32
36
40
44
48
52
56
OB-302
% Change
0
-5
-10
-15
-20
8 28 48
Week

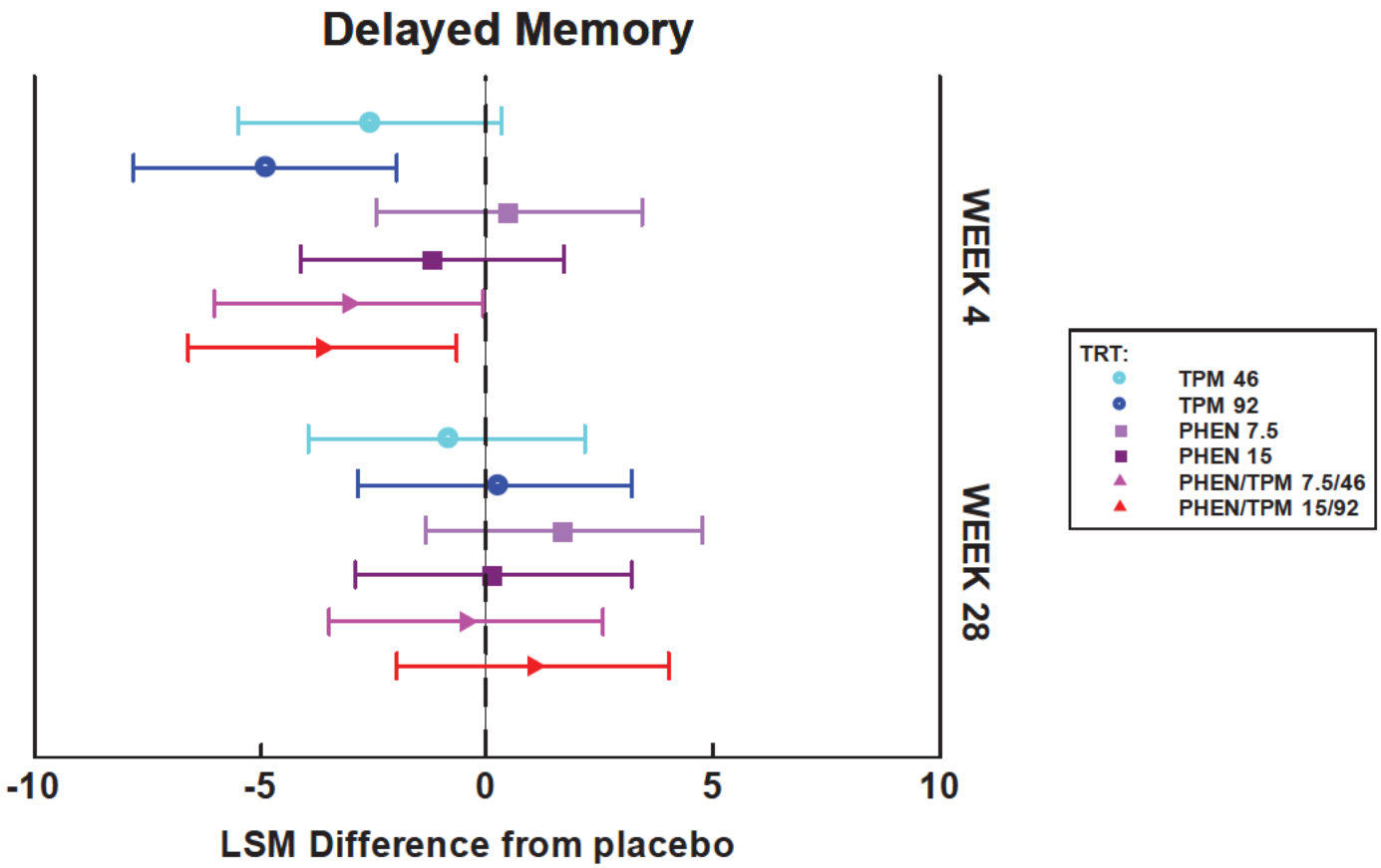

OB-303
Boxplots for PCHG
20
0
-5
-10
-20
-40
8: Placebo
8: 7.5Ph/46T
8: 15Ph/92T
12: Placebo
12: 7.5Ph/46T
12: 15Ph/92T
16: Placebo
16: 7.5Ph/46T
16: 15Ph/92T
20: Placebo
20: 7.5Ph/46T
20: 15Ph/92T
24: Placebo
24: 7.5Ph/46T
24: 15Ph/92T
28: Placebo
28: 7.5Ph/46T
28: 15Ph/92T
32: Placebo
32: 7.5Ph/46T
32: 15Ph/92T
36: Placebo
36: 7.5Ph/46T
36: 15Ph/92T
40: Placebo
40: 7.5Ph/46T
40: 15Ph/92T
44: Placebo
44: 7.5Ph/46T
44: 15Ph/92T
48: Placebo
48: 7.5Ph/46T
48: 15Ph/92T
52: Placebo
52: 7.5Ph/46T
52: 15Ph/92T
56: Placebo
56: 7.5Ph/46T
56: 15Ph/92T
Cohort of patients at Week: TRT

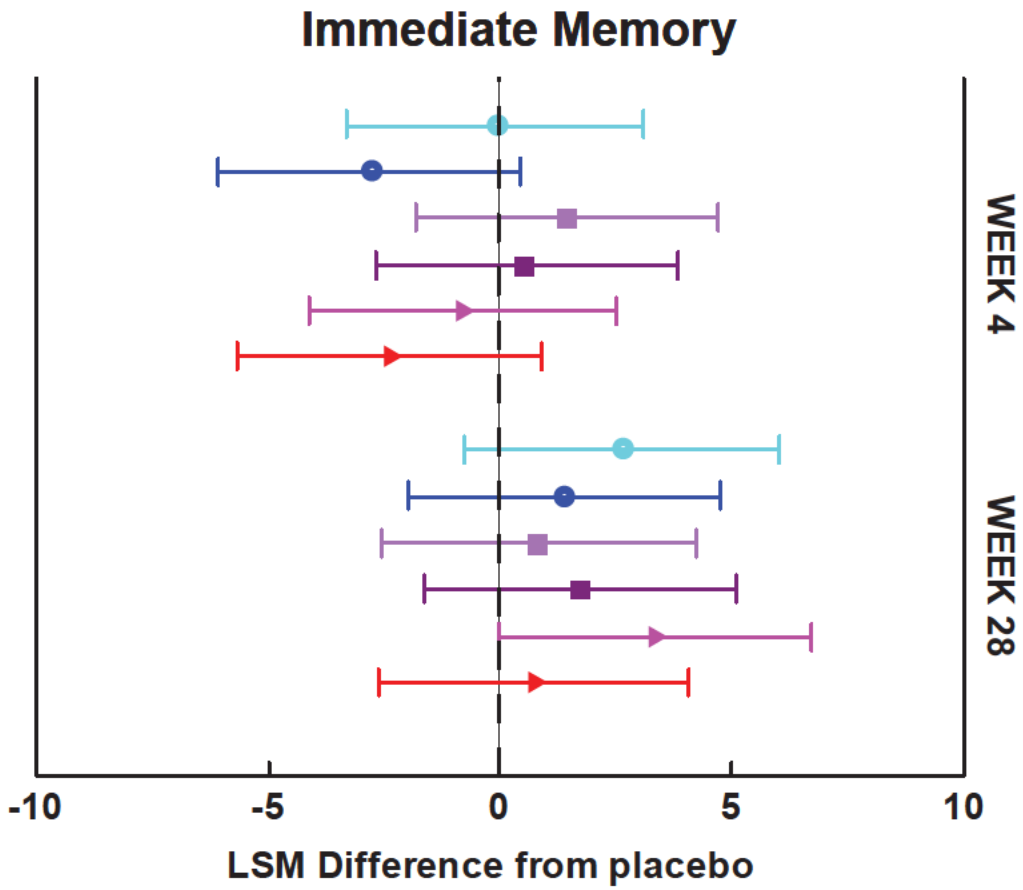

Cohort of patients at Week
4
8
12
16
20
24
28
32
36
40
44
48
52
56
OB-303
% Change
0
-5
-10
-15
-20
8 28 48
Week

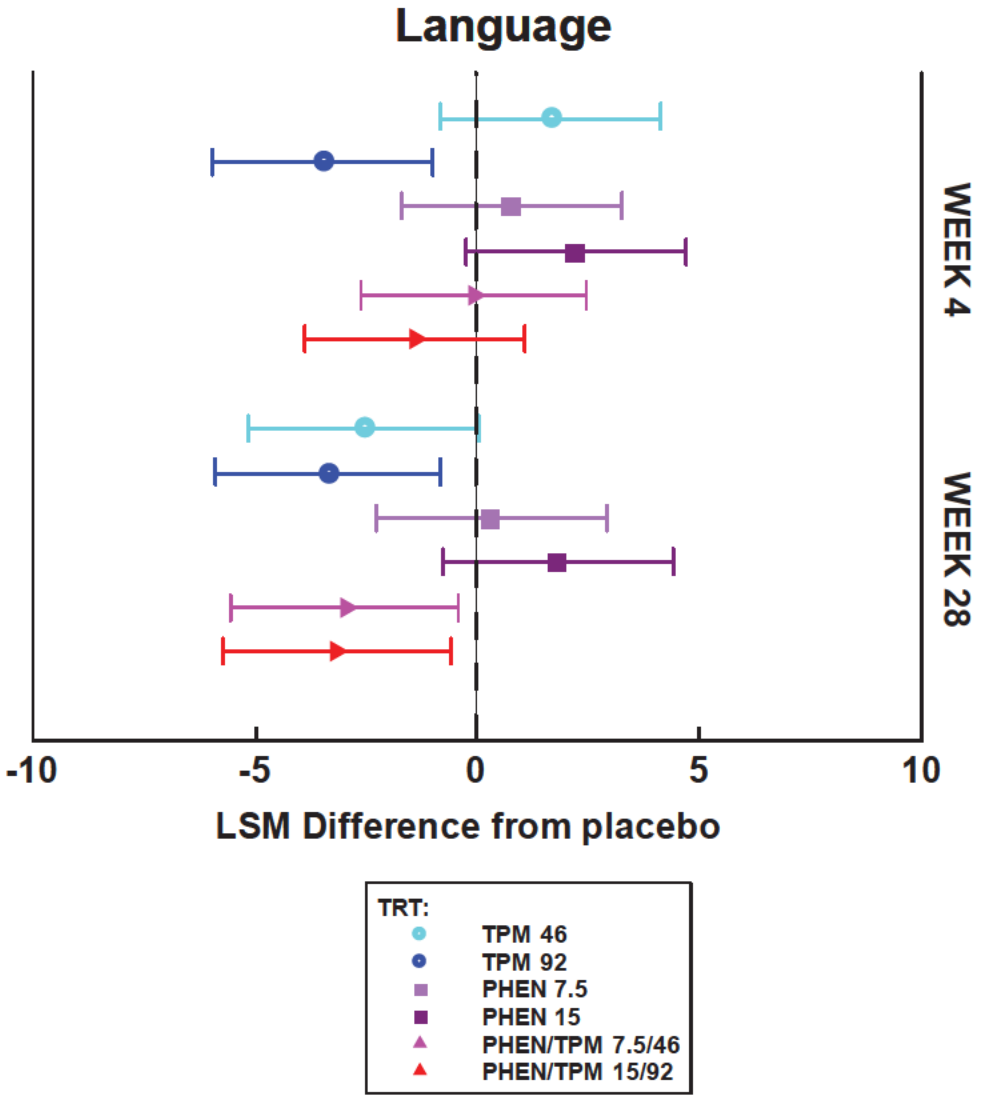
4. RBANS (Repeatable Battery for the Assessment of Neuropsychological Status)

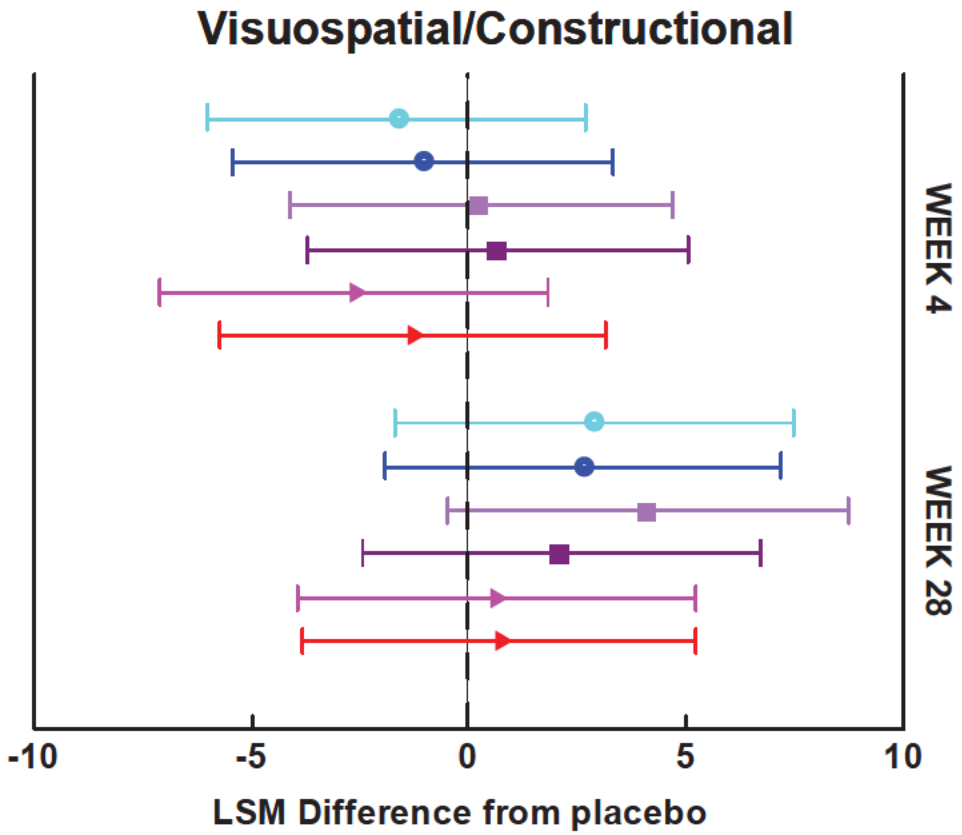
RBANS is a brief, individually administered test measuring attention, language, visuospatial/constructional abilities, and immediate and delayed memory. The following figures display the LSM treatment differences (95% CI) from placebo for individual domain and total scale of index scores. A negative change means worse.

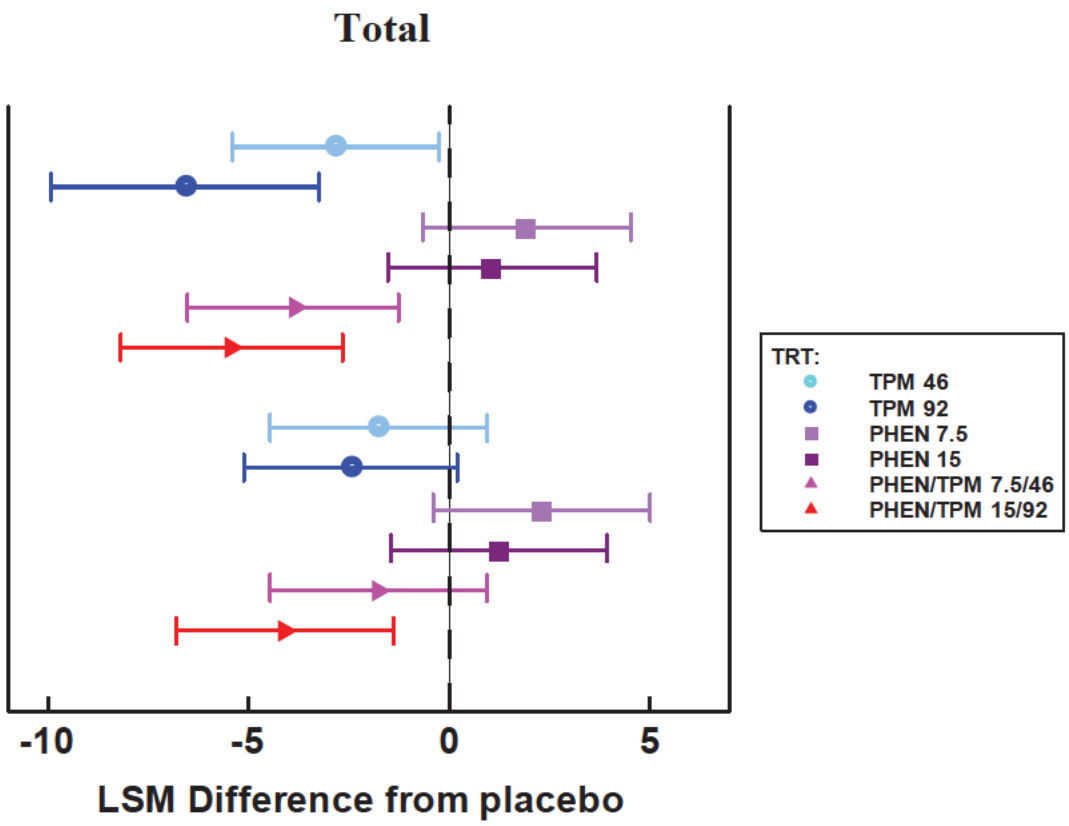



Delayed Memory
WEEK 4
WEEK 28
-10
-5
0
5
10
LSM Difference from placebo
TRT:
TPM 46
TPM 92
PHEN 7.5
PHEN 15
PHEN/TPM 7.5/46
PHEN/TPM 15/92


Immediate Memory
WEEK 4
WEEK 28
-10
-5
0
5
10
LSM Difference from placebo
TRT:
TPM 46
TPM 92
PHEN 7.5
PHEN 15
PHEN/TPM 7.5/46
PHEN/TPM 15/92


Language
WEEK 4
WEEK 28
-10
-5
0
5
10
LSM Difference from placebo
TRT:
TPM 46
TPM 92
PHEN 7.5
PHEN 15
PHEN/TPM 7.5/46
PHEN/TPM 15/92


Visuospatial/Constructional
WEEK 4
WEEK 28
-10
-5
0
5
10
LSM Difference from placebo
TRT:
TPM 46
TPM 92
PHEN 7.5
PHEN 15
PHEN/TPM 7.5/46
PHEN/TPM 15/92


Total
TRT:
TPM 46
TPM 92
PHEN 7.5
PHEN 15
PHEN/TPM 7.5/46
PHEN/TPM 15/92
-10
-5
0
5
LSM Difference from placebo

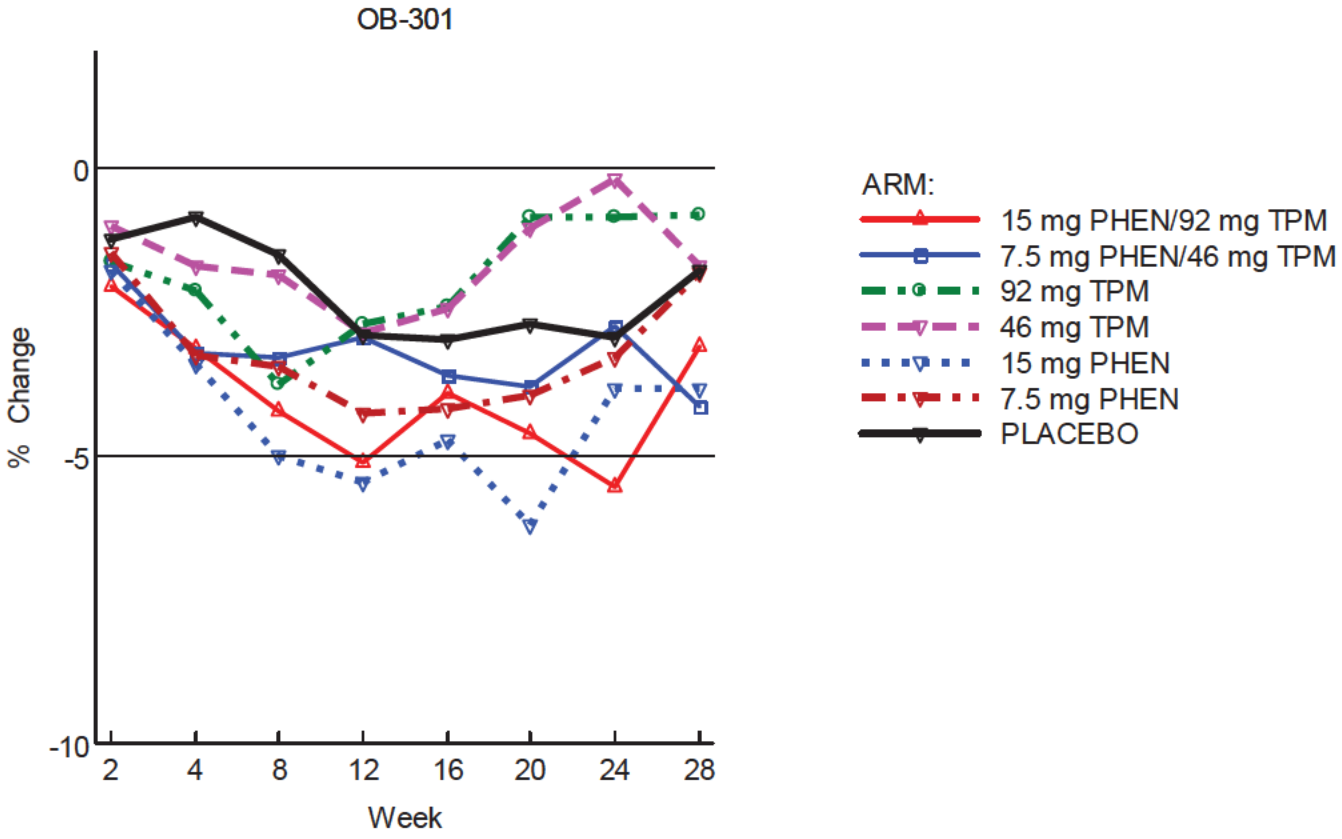
APPEARS THIS WAY ON ORIGINAL

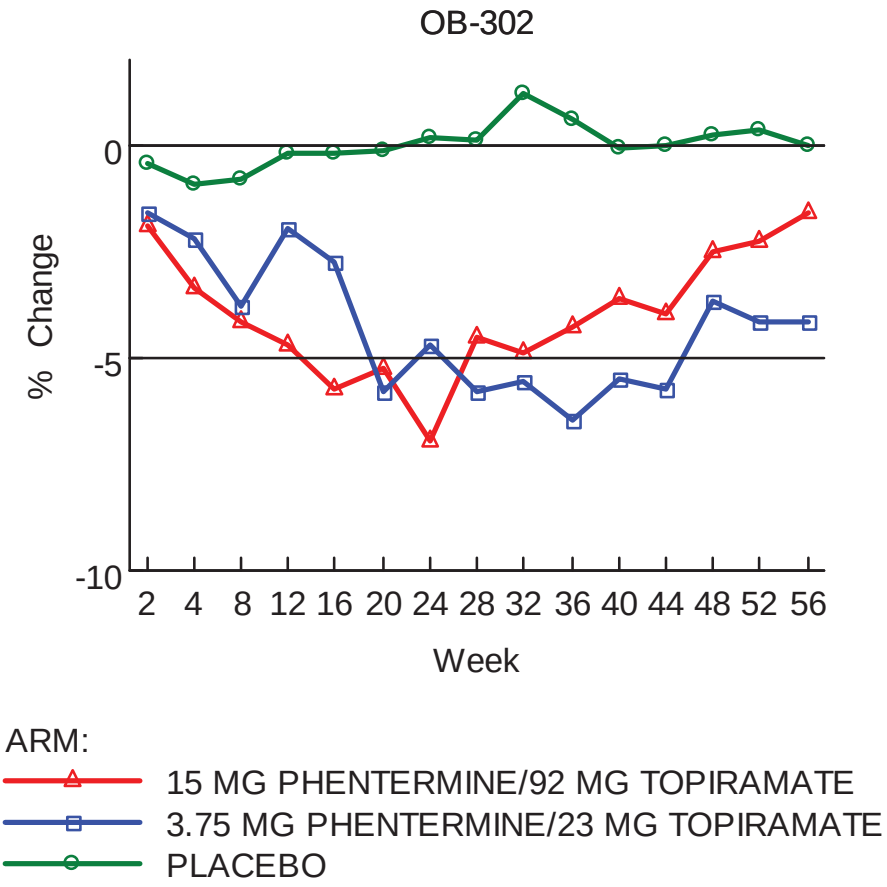
5. On study patients who stopped treatment

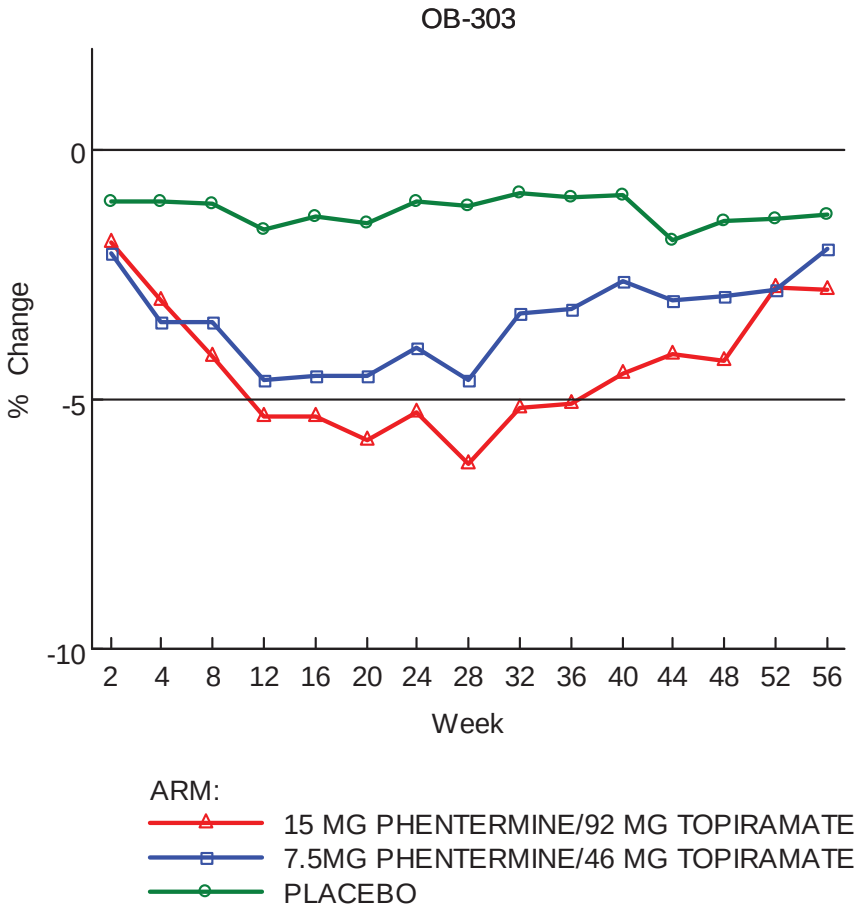
% of Patients Who Stopped Treatment, not Study Visits

Study	15 PHEN/92 TPM	15 PHEN	92 TPM	7.5 PHEN/46 TPM	7.5 PHEN	46 TPM	3.75 PHTN/23 TPM	Placebo
301	7/103 (7%)	8/106 (8%)	10/105 (10%)	5/103 (5%)	5/104 (5%)	6/102 (6%)		5/103 (5%)
302	39/498 (8%)						9/234 (4%)	31/257 (6%)
303	99/981 (10%)			30/488 (6%)				52/979 (5%)

Median % Weight Change from Baseline for Patients Who Stopped Treatment, not Study Visits




OB-302
0
-5
-10
% Change
2 4 8 12 16 20 24 28 32 36 40 44 48 52 56
Week
ARM:
15 MG PHENTERMINE/92 MG TOPIRAMATE
3.75 MG PHENTERMINE/23 MG TOPIRAMATE
PLACEBO


OB-303
0
-5
-10
% Change
2
4
8
12
16
20
24
28
32
36
40
44
48
52
56
Week
ARM:
15 MG PHENTERMINE/92 MG TOPIRAMATE
7.5MG PHENTERMINE/46 MG TOPIRAMATE
PLACEBO

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

/s/

LEE PING PIAN
09/27/2010

JON T SAHLROOT
09/27/2010
concur

US 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

New Drug Application

Clinical Studies

NDA/Serial Number:	22-580/SN000
Drug Name:	Qnexa (phentermine and topiramate)
Indication(s):	Treatment of obesity in patients with and without weight-related comorbidities
Applicant:	Vivus, Inc.
Dates:	Submitted: 12/28/2009 PDUFA: 10/28/2010
Review Priority:	Standard
Biometrics Division:	Division of Biometrics VII
Statistics Reviewer:	Benjamin Neustifter, Ph.D.
Concurring Reviewers:	Mat Soukup, Ph.D. Aloka Chakravarty, Ph.D.
Medical Division:	Division of Metabolism and Endocrinology Products (DMEP)
Clinical Team:	Reviewer: Mary Roberts, M.D. Lead: Eric Colman, M.D.
Project Manager:	Pooja Dharia (DMEP)
Keywords:	combination product, safety, suicidality

Contents

1 Executive Summary **4**

1.1 Conclusions and Recommendations 4

1.2 Brief Overview of Clinical Studies 4

1.3 Statistical Issues and Findings 5

2 Introduction **5**

2.1 Product Description 5

2.2 Clinical Trial Overview 5

3 Evaluation of Safety Regarding Depression and Suicidality in the Phase 3 Studies **6**

3.1 Statistical Methodology 6

3.2 Study Demographics 7

3.3 C-SSRS Description 7

3.3.1 Suicidal Ideations 8

3.3.2 Suicidal Behaviors 9

3.4 Primary Outcomes 9

3.4.1 Ideations 9

3.4.2 Behaviors 12

3.5 Secondary Outcomes 12

3.5.1 Ideations 12

3.5.2 Behaviors 13

3.6 Increased Symptoms from Baseline 13

4 Summary and Conclusions **14**

4.1 Conclusions and Recommendations 15

Appendix **16**

A.1 Supplementary Tables 16

A.1.1 Treatment Arm Demographics 16

A.1.2 Comparison of Qnexa 3.75/23 mg to Placebo 19

A.1.3 Post-Baseline Ideation Severity for Most Severe Ideation 20

A.1.4 Post-Baseline Controllability for Most Severe Ideation 21

A.1.5 Post-Baseline Deterrents for Most Severe Ideation 22

A.1.6 Post-Baseline Duration for Most Severe Ideation 23

A.1.7 Post-Baseline Frequency for Most Severe Ideation 24

A.1.8 Post-Baseline Reasons for Most Severe Ideation 25

A.1.9 Type of Worsening Suicidality by Treatment Arm 26

Signatures/Distribution List **27**

1 Executive Summary

1.1 Conclusions and Recommendations

Qnexa is a fixed-dose combination of immediate-release phentermine (PHEN) and modified-release topiramate (TPM), indicated for the treatment of obesity. The sponsor, Vivus, Inc., performed a series of studies to determine whether the weight loss related to treatment with Qnexa is greater than that of placebo treatment, as well as active comparator treatments of single-agent PHEN and TPM. The Division of Metabolism and Endocrinology Products (DMEP) requested a statistical consult to analyze clinical trial data from three Phase 3 studies to explore for possible effects of Qnexa on depression and suicidality.

The three studies examined were powered for efficacy and not for safety (e.g. depression and suicidality). Further, the number of events (suicidal ideations and/or behaviors) in the provided studies is low, making it difficult to draw a definite conclusion in regards to differences between the proposed drug and a placebo treatment. Based on evaluation of the safety data regarding the Columbia – Suicide Severity Rating Scale (C-SSRS), the following conclusions and recommendations are offered:

1. From the data available, there does not appear to be a significant relationship between any dose of Qnexa and an increase in suicidal ideations or behaviors above a placebo treatment.

2. There may be a mild relationship between the high-dose (15/92 mg) Qnexa and increased severity of suicidal ideations in the categories of controllability, duration, and frequency above a placebo treatment. The possible increase appears to be fairly mild, if it exists, and it cannot be determined whether it is significant without further data from appropriately designed studies.

3. There does not appear to be a significant relationship between the proposed drug and an emergence or worsening in suicidal ideations or behaviors in subjects.

1.2 Brief Overview of Clinical Studies

The three studies, OB-301, OB-302, and OB-303, were all randomized, double-blind, placebo-controlled clinical trials carried out from November 2007 to June 2009. Each of the studies was individually powered to provide 90% or greater power for the detection of efficacy difference between the mid-dose (7.5/46 mg) Qnexa treatment and a placebo. These studies contained varying dosages of Qnexa, PHEN, and TPM, although all contained placebo and Qnexa 15/92mg treatment arms. Since Qnexa is a combination product, its first dosage amount refers to the amount of PHEN it contains, and the second dosage amount to the amount of TPM. Phase 2

studies were also conducted, but were not considered in the data exploration and analysis in this report since they did not include data from the C-SSRS on suicidality.

1.3 Statistical Issues and Findings

As the studies under consideration were designed to assess efficacy (as measured primarily by percent weight loss), with safety as a secondary endpoint, the sample sizes were not chosen to power safety-related statistical infererence. Suicidal behaviors, and to a lesser extent suicidal ideations, are rare events, and so the sample sizes for the Phase 3 studies do not contain enough events to allow for the detection of meaningful differences. Exploratory data analysis, however, was performed in order to suggest possible safety signals regarding suicidality for further study in controlled clinical trials. The majority of data exploration showed no apparent differences between any of the doses of Qnexa and a placebo treatment with respect to suicidality; however, a mild possible relationship between the 15/92 mg dose of Qnexa and an increase in some severity subscales for suicidal ideations above placebo was noted.

2 Introduction

2.1 Product Description

Qnexa is a fixed-dose combination of phentermine (PHEN) and topiramate (TPM). The sponsor states that Qnexa is manufactured in four dosages (PHEN/TPM): 3.75/23 mg, 7.5/46 mg, 11.25/69 mg, and 15/92 mg, with 7.5/46 mg as the sponsor-proposed recommended dosage. The sponsor suggests that the 15/92 mg dose is proposed for subjects not responding to the 7.5/46 mg dose, and the remaining dosages are recommended for titration as needed. Qnexa is proposed for the treatment of obesity; PHEN was previously approved by the FDA for appetite suppression and treatment of obesity, and TPM was previously approved as an anticonvulsive and for treatment of epilepsy and migraines. Qnexa is provided in capsules to be taken once daily.

2.2 Clinical Trial Overview

The Phase 3 program consisted of three double-blind, randomized, placebo-controlled clinical trials. Each trial contained an arm with 15/92 mg dose Qnexa, and an arm with a placebo; doses of single-agent PHEN and TPM and alternate doses of Qnexa varied by study. Table 1 provides the sample sizes, arms, and dates for the studies.

Table 1: Study overviews

Study	Arm	Sample Size (Ignoring Screen Failures)	Dates (From first enrollment to last subject completion)
OB-301	Placebo	109	12-3-2007 – 9-30-2008
	PHEN 7.5 mg	109	
	PHEN 15 mg	108	
	TPM 46 mg	108	
	TPM 92 mg	107	
	Qnexa 7.5/46 mg	107	
	Qnexa 15/92 mg	108	
	Total Sample	756	
OB-302	Placebo	514	11-1-2007 – 5-19-2009
	Qnexa 3.75/23 mg	241	
	Qnexa 15/92 mg	512	
	Total Sample	1267	
OB-303	Placebo	994	11-1-2007 – 6-30-2009
	Qnexa 7.5/46 mg	498	
	Qnexa 15/92 mg	995	
	Total Sample	2487	

Source: OB-301, OB-302, OB-303 Study Reports, Table 14.1.2. Recreated by reviewer.

3 Evaluation of Safety Regarding Depression and Suicidality in the Phase 3 Studies

3.1 Statistical Methodology

The three Phase 3 studies were designed with efficacy as a primary endpoint, as measured by percent weight loss and percentage of subjects with at least 5% weight loss by the end of the treatment period. While safety measures were included, the studies were not powered to detect safety signals, nor otherwise constructed with safety as a primary endpoint. For this reason, standard statistical inference were not included in the protocols for these studies. Exploratory data analyses, however, were used to inspect the data for possible safety signals.

Tabulations according to the scoring guide for the Columbia – Suicide Severity Rating Scale (C-SSRS), described in Section 3.3 were used to compare the treatment arms with regards to suicidal ideations and behaviors. Odds Ratios (ORs) and the associated 95% Confidence Intervals (using the delta-method) for the risk of having types of suicidal ideations were calculated to demonstrate the lack of strong apparent relationship between treatments. Subjects who ap-

Table 2: Study Demographics, N (% of Study)

		OB-301 (N=756)	OB-302 (N=217)	OB-303 (N=2487)
Sex	Female	599 (79.2)	1050 (82.9)	1737 (69.8)
	Male	157 (20.8)	217 (17.1)	750 (30.2)
Race	American Indian or Alaska Native	4 (0.5)	10 (0.8)	12 (0.5)
	Asian	6 (0.8)	5 (0.4)	18 (0.7)
	Black or African-American	140 (18.5)	218 (17.2)	295 (11.5)
	Native Hawaiian or other Pacific Islander	1 (0.1)	4 (0.3)	3 (0.2)
	White	594 (78.6)	1003 (79.2)	2130 (85.7)
	Multiple	6 (0.8)	12 (1.0)	14 (0.6)
	Other	5 (0.7)	15 (1.2)	24 (1.0)
Ethnicity	Hispanic or Latino	68 (9.0)	184 (14.5)	328 (13.2)
	Not Hispanic/Latino	688 (91.0)	1083 (85.5)	2159 (86.8)
Age	Mean (Std. Dev.)	45.1 (11.9)	42 (11.8)	50.6 (10.4)
Base Weight	Mean (Std. Dev.)	101.3 (15.5)	116.1 (21.2)	103.0 (17.9)

Source: Table 14.1.4 from OB-301, OB-302, OB-303 Study Reports. Recreated by reviewer.

peared to demonstrate worsening or emergent suicidal ideations and behaviors had their data flagged and further analyzed, as detailed in section 3.6.

3.2 Study Demographics

All three of the studies were performed in the United States of America, and tended to enroll similar subject populations in which a higher percentage of subjects were white (78.6%, 79.2%, and 85.7% of enrolled subjects for studies 301, 302, and 303 respectively) and female (79.2%, 82.9%, 69.8%), with an mean subject age in the range of 40–50 and mean baseline weight in the range of 100–116 kg. Table 2 summarizes the demographics for subjects in each study. In the Appendix, Tables 9 through 11 summarize the demographics by treatment arm.

3.3 C-SSRS Description

For the analysis of depression and suicidality, primary data were taken from the C-SSRS. The C-SSRS is an 11-item clinician-administered questionnaire that is made to measure both suicidal

ideations and behavior. According to the study protocols, this assessment was to be given to all subjects at the screening visit for a baseline value, and then subsequently re-administered every 2–4 weeks at Visits 3 (Week 2) through 10 (Week 28) if and only if the subject met certain criteria: scoring above a prespecified depression level on the PHQ-9 instrument; answering certain questions regarding suicidal ideations in the affirmative; or experiencing adverse events that may be potentially related to suicidal ideation.

3.3.1 Suicidal Ideations

The C-SSRS contains 5 Yes/No questions regarding different levels of severity for suicidal ideations, asking if the subject:

- Wished to be dead
- Had any thoughts about killing herself/himself
- Had considered the methods for a suicide attempt
- Had some intention of acting on these plans
- Had begun or completed a plan to kill herself/himself

Any question answered "Yes" had a space for the clinician to record the specifics of the subject's ideations. If the subject answered "Yes" to any of the suicidal ideation questions above, the clinician was then supposed to answer a further series of questions regarding the severity of the ideations. Severity of an ideation was rated on a scale from 1 to 5, with 1 corresponding to the "wish to be dead" question above, and 5 to the subject having begun or completed a plan to kill herself/himself. The severity subscales/questions related to:

- Frequency of ideations
- Duration of ideations
- Controllability of ideations
- Deterrents that conflicted with ideations
- Reasons for ideations

Each of these questions was also ranked on a severity scale of 1 (least severe) to 5 (most severe), with some having a 0 ranking (for "does not apply") also allowable. Clinicians were supposed to collect these values for the subject's "Most Common" and "Most Severe" ideations during the previous time period, if the subject had any ideations. Any subject that did not have an ideation during the period was not given the questionnaire and recorded as answering 0 to all questions.

3.3.2 Suicidal Behaviors

The assessment also includes a series of Yes/No questions (with follow-ups for Yes answers) regarding suicidal behaviors; these questions are administered regardless of the subject's answers to the suicidal ideation questions. The questions ask if the subject has ever committed:

- A suicide attempt
- An interrupted suicide attempt
- An aborted suicide attempt
- Preparatory acts or behaviors
- Suicidal behavior. Note that this question is not explained in either the questionnaire or study protocol. It is not a composite variable for the other suicidal behavior questions, as some subjects answered this question "Yes" with all the others answered "'No." It is unclear what definition of "Suicidal Behavior" is used for the application of this question.

For attempted suicides, the clinician also was required to answer questions regarding the actual lethality/medical damage of the attempt, and regarding the potential lethality of the act if no actual lethality occurred.

3.4 Primary Outcomes

3.4.1 Ideations

Table 3 gives, for the Placebo and Qnexa arms, the number of subjects across studies that reported each type of suicidal ideation at a visit other than the baseline. The final column gives a composite total of all unique subjects that reported any ideation at a visit other than baseline. There does not appear to be any significant relationship between Qnexa use at any dose and an increase in suicidal ideations from these data.

While little difference is observed between the Qnexa and placebo treatments, Tables 4 and 5 provide the odds ratios (ORs) and associated 95% confidence intervals for the two primary Qnexa doses (7.5/46 mg and 15/92 mg) versus placebo treatment for suicidal ideations. Since there were so few suicidal behaviors observed during the studies, risk comparisons result in too large confidence intervals to draw any inference about an increase in risk. However, Tables 4 and 5 provide the OR and confidence intervals for the composite total of subjects on each treatment that had at least one ideation and/or behavior. Table 6 provides the same calculations for all of the Qnexa treatments combined against placebo. A comparison of the low-dose (3.75/23 mg) Qnexa to placebo is included in the Appendix in Table 12. All calculations are based solely on post-baseline events. Note that the Relative Risks (RRs) and associated confidence intervals

Table 3: Ideations post-baseline, N of subjects experiencing (% of arm)

	Placebo (N=1617)	Qnexa 3.75/23 mg (N=241)	Qnexa 7.5/46 mg (N=605)	Qnexa 15/92 mg (N=1615)	All Qnexa (N=2461)
Wish to be Dead	9 (0.56)	1 (0.41)	3 (0.50)	15 (0.93)	19 (0.77)
Non-Specific Active Suicidal Thoughts	5 (0.31)	1 (0.41)	1 (0.17)	6 (0.37)	8 (0.33)
Active Ideation: Methods, No Intent	2 (0.12)	0 (0)	0 (0)	1 (0.06)	1 (0.04)
Active Ideation: Intent, No Plan	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Active Ideation: Plan & Intent	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Composite Total (Unique subjects)	11 (0.68)	1 (0.41)	3 (0.50)	16 (0.99)	20 (0.81)

Source: Created by reviewer

Table 4: OR for Subjects Experiencing Post-Baseline Event: Qnexa 7.5/46 mg against Placebo

	Qnexa 7.5/46 (N=605) N (%)	Placebo (N=1617) N (%)	Odds Ratio (95% CI)
Wish to be Dead	3 (0.50)	9 (0.56)	0.89 (0.24, 3.30)
Non-Specific Active Suicidal Thoughts	1 (0.17)	5 (0.31)	0.53 (0.06, 4.58)
Active Ideation: Methods, No Intent	0 (0)	2 (0.12)	*
Active Ideation: Intent, No Plan	0 (0)	0 (0)	*
Active Ideation: Plan & Intent	0 (0)	0 (0)	*
Composite Total (by Unique Subjects)	3 (0.50)	11 (0.68)	0.73 (0.20, 2.62)
Composite Total with Behaviors	3 (0.50)	12 (0.74)	0.67 (0.19, 2.37)

* Either arm having 0 subjects with events means OR cannot be calculated

Source: Calculated by reviewer

Table 5: OR for Subjects Experiencing Post-Baseline Event: Qnexa 15/92 mg against Placebo

	Qnexa 15/92 (N=1615) N (%)	Placebo (N=1617) N (%)	Odds Ratio (95% CI)
Wish to be Dead	15 (0.93)	9 (0.56)	1.68 (0.73, 3.84)
Non-Specific Active Suicidal Thoughts	6 (0.37)	5 (0.31)	1.20 (0.37, 3.95)
Active Ideation: Methods, No Intent	1 (0.06)	2 (0.12)	0.50 (0.05, 5.52)
Active Ideation: Intent, No Plan	0 (0)	0 (0)	*
Active Ideation: Plan & Intent	0 (0)	0 (0)	*
Composite Total (by Unique Subjects)	16 (0.99)	11 (0.68)	1.46 (0.68, 3.16)
Composite Total with Behaviors	17 (1.05)	12 (0.74)	1.42 (0.68, 2.99)

* Either arm having 0 subjects with events means OR cannot be calculated

Source: Calculated by reviewer

Table 6: OR for Subjects Experiencing Post-Baseline Event: All Qnexa against Placebo

	All Qnexa (N=2461) N (%)	Placebo (N=1617) N (%)	Odds Ratio (95% CI)
Wish to be Dead	19 (0.77)	9 (0.56)	1.39 (0.63, 3.08)
Non-Specific Active Suicidal Thoughts	8 (0.33)	5 (0.31)	1.05 (0.34, 3.22)
Active Ideation: Methods, No Intent	1 (0.04)	2 (0.12)	0.33 (0.03, 3.62)
Active Ideation: Intent, No Plan	0 (0)	0 (0)	*
Active Ideation: Plan & Intent	0 (0)	0 (0)	*
Composite Total (by Unique Subjects)	20 (0.81)	11 (0.68)	1.20 (0.57, 2.50)
Composite Total with Behaviors	21 (0.85)	12 (0.74)	1.15 (0.56, 2.34)

* Either arm having 0 subjects with events means OR cannot be calculated

Source: Calculated by reviewer

will be nearly identical for each of these comparisons, due to the low event rate, and so are not included in these tables.

There does not appear to be any obvious safety signal that indicates a significant increase in risk of any suicidal ideations from placebo to either of the higher Qnexa doses. It is important to reiterate, however, that the studies were not powered to detect these safety endpoings. Also, it is worth noting that while the OR confidence intervals all included 1, indicating no significant increase in risk of ideations for the Qnexa arm above the placebo, the ORs for the 15/92 mg dose of Qnexa over placebo are larger than those for the 7.5/46 mg dose.

3.4.2 Behaviors

Similarly to Table 3, Table 7 shows the number of subjects that experienced any of the types of suicidal behavior across studies. Note that there were only 3 subjects who experienced events between the Placebo and Qnexa arms: 2 subjects on the Placebo treatment experienced suicidal events (one attempted suicide and had suicidal behavior, the other only demonstrated suicidal behavior), and 1 subject on the 15/92 mg dose Qnexa had suicidal behavior.

3.5 Secondary Outcomes

3.5.1 Ideations

As discussed in Section 3.3, the C-SSRS contains questions regarding the severity of suicidal ideations. While the event sizes are small, they are provided to give possible information on relationships that may exist between use of Qnexa treatment and an increase in ideation severity. In the Appendix, Tables 13 through 18 give the severity level and each of the severity subscale measures (commonality, deterrents, duration, frequency, reasons) for the "most severe" suicidal ideations experienced by subjects on the placebo or Qnexa treatments. These Tables were constructed to contain only one record per subject; thus, any subject with more than one ideation was taken as the highest of the recorded severity or severity subscale ratings. The results from subjects responses for their "most common" suicidal ideation were not significantly different from those provided in the Tables. Note that in one or two subjects' cases, the severity subscales were not administered, for unknown reasons, resulting in a discepancy in total numbers from Table 3.

There does not appear to be any strong relationship between the proposed drug and an increase in recorded ideation severity over a placebo. However, while it is not possible to quantify its significance due to the nature of the studies and the sample sizes, there may be a mild signal in some of the severity subscales for the Qnexa 15/92 mg dose. Specifically, in considering Tables 14, 16, and 17, it appears that suicidal ideations in the high-dose Qnexa arm may tend toward more severe reactions on these subscales than the placebo treatment; that

Table 7: Behaviors post-baseline, N of subjects experiencing (% of arm)

	Placebo (N=1617)	Qnexa 3.75/23 mg (N=241)	Qnexa 7.5/46 mg (N=605)	Qnexa 15/92 mg (N=1615)	All Qnexa (N=2461)
Suicide Attempt	1 (0.06)	0 (0)	0 (0)	0 (0)	0 (0)
Interrupted Attempt	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Aborted Attempt	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Preparatory Behaviors	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Suicidal Behavior	2 (0.12)	0 (0)	0 (0)	1 (0.06)	1 (0.04)
Composite Total (Unique subjects)	2 (0.12)	0 (0)	0 (0)	1 (0.06)	1 (0.13)

Source: Created by reviewer

is, the "most severe" ideations on the 15/92 mg Qnexa were reported to be harder to control, lasted longer, and/or were more frequent that the "most severe" ideations on the placebo.

3.5.2 Behaviors

The C-SSRS also contained questions regarding the lethality of suicide attempts; these were only administered if a subject had showed some form of suicide attempt (not if the subject had only demonstrated "suicidal behavior"). Only 2 suicide attempts were recorded during the assessment period. One attempt of Lethality 0, "No physical damage or very minor physical damage (e.g. surface scratches)," was recorded for a subject in the Placebo study arm, and one attempt of Lethality 3, "Moderately severe physical damage; medical attention needed (e.g. conscious but sleepy, somewhat responsive; second-degree burns; bleeding of major vessel)," was recorded for a subject in the 15 mg PHEN arm. Since there were so few recorded suicide attempts, statistical inference of these data are not possible.

3.6 Increased Symptoms from Baseline

Previous results captured post-baseline events, but did not take into account the baseline measurements for subjects, and so did not account for possible histories of suicidality. To account for this, the data were than analyzed in order to isolate cases where subjects appeared to have more severe suicidal symptoms than at their baseline analysis. In this analysis, subjects were considered to have worsened, with regard to suicidal ideations and behaviors, if any of the following occurred:

- Subject had any type of suicidal ideations or behaviors when she/he did not have any at the baseline assessment
- Subject had more severe suicidal ideations than she/he had at the baseline assessment
- Subject recorded responses to any of the other ideation severity categories (controllability, frequency, deterrents, duration, reasons) that were higher than at baseline
- Subject demonstrated higher lethality suicidal behaviors than at baseline

Only subjects that were recorded as being on-treatment, or off-treatment for less than 28 days[1], were considered to have worsened in a way that may be due to treatment.

Table 8 shows the number of subjects considered to have worsened at some point during the study, according to the above criteria.

There does not appear to be any particularly strong treatment effect in terms of patients with worsening suicidal ideations or behaviors for the Qnexa arms of the trial; the proportion of subjects experiencing such worsening appears similar to the placebo arm. Table 19, which further gives the general category of the subject's worsening (e.g. emergence of suicidal ideations, emergence of suicidal behaviors, increased severity of ideations) by arm, is included in the Appendix. As above, there appears to be little noticeable difference between the Qnexa treatments and placebo.

4 Summary and Conclusions

The three Phase 3 studies used to determine the efficacy of Qnexa were not powered for detecting differences in safety endpoints. Therefore, the data analysis carried out in this memorandum were exploratory in nature. Exploration of data collected using the Columbia – Suicide Severity Rating Scale (C-SSRS) does not appear to show any strong relationship between use of the proposed drug and an increase in suicidal ideations and/or behaviors above a placebo treatment.

Inspection of results from the ideation severity subscales on the C-SSRS shows data that may suggest a slight relationship between use of the high-dose (15/92 mg) Qnexa and an increase in the severity of suicidal ideations above placebo with regard to their controllability, duration, and frequency: see Tables 14, 16, and 17.

Finally, the data were specifically explored to find subjects that had emergent or worsened suicidal ideations and/or behaviors from their baseline measurements; as with the number of ideations and behaviors, there did not appear to be any relationship between the use of Qnexa and the number of subjects who had emergent or worsened suicidal symptoms.

[1]The cutoff of 28 days as the point at which treatment was no longer relevant was determined based on input from the clinical reviewer.

Table 8: Subjects that Worsened from Baseline, N (% of Arm)

	Placebo (N=1617)	Qnexa 3.75/23 mg (N=241)	Qnexa 7.5/46 mg (N=605)	Qnexa 15/92 mg (N=1615)	All Qnexa (N=2461)
Worsened Within 28 Days of Trt	11 (0.68)	1 (0.41)	2 (0.33)	11 (0.68)	14 (0.57)

Source: Calculated by reviewer.

4.1 Conclusions and Recommendations

The three studies examined were powered for efficacy and not for safety (e.g. depression and suicidality). Further, the number of events (suicidal ideations and/or behaviors) in the provided studies is low, making it difficult to draw a definite conclusion in regards to differences between the proposed drug and a placebo treatment. Based on evaluation of the safety data regarding the Columbia – Suicide Severity Rating Scale (C-SSRS), the following conclusions and recommendations are offered:

1. From the data available, there does not appear to be a significant relationship between any dose of Qnexa and an increase in suicidal ideations or behaviors above a placebo treatment.

2. There may be a mild relationship between the high-dose (15/92 mg) Qnexa and increased severity of suicidal ideations in the categories of controllability, duration, and frequency above a placebo treatment. The possible increase appears to be fairly mild, if it exists, and it cannot be determined whether it is significant without further data from appropriately designed studies.

3. There does not appear to be a significant relationship between the proposed drug and an emergence or worsening in suicidal ideations or behaviors in subjects.

As the studies under consideration were designed to assess efficacy (as measured primarily by percent weight loss), with safety as a secondary endpoint, the sample sizes were not chosen to power safety-related statistical infererence. Suicidal behaviors, and to a lesser extent suicidal ideations, are rare events, and so the sample sizes for the Phase 3 studies do not contain enough events to allow for the detection of meaningful differences. Exploratory data analysis, however, was performed in order to suggest possible safety signals regarding suicidality for further study in controlled clinical trials. The majority of data exploration showed no apparent differences between any of the doses of Qnexa and a placebo treatment with respect to suicidality; however, a mild possible relationship between the 15/92 mg dose of Qnexa and an increase in some severity subscales for suicidal ideations above placebo was noted.

Appendix

A.1 Supplementary Tables

A.1.1 Treatment Arm Demographics

Table 9: Arm Demographics, N (% of Arm)

		Placebo (N=1617)	7.5 mg PHEN (N=109)	15 mg PHEN (N=108)
Sex	Female	1206 (74.6)	434 (71.7)	86 (79.6)
	Male	411 (9.1)	23 (21.1)	22 (20.4)
Race	American Indian or Alaska Native	7 (0.4)	0 (0)	1 (0.9)
	Asian	8 (0.5)	1 (0.9)	1 (0.9)
	Black or African-American	225 (13.9)	26 (23.9)	14 (13.0)
	Native Hawaiian or other Pacific Islander	2 (0.1)	0 (0)	1 (0.9)
	White	1349 (83.4)	79 (72.5)	90 (83.3)
	Multiple	11 (0.7)	2 (1.8)	0 (0)
	Other	15 (0.9)	1 (0.9)	1 (0.9)
Ethnicity	Hispanic or Latino	215 (13.3)	9 (8.3)	9 (8.3)
	Not Hispanic/Latino	1402 (86.7)	100 (91.7)	99 (91.7)
Age	Mean (Std. Dev.)	47.7 (11.5)	45.9 (11.6)	54.2 (12.3)
Base Weight	Mean (Std. Dev.)	107.1 (19.9)	101.0 (15.1)	101.3 (16.4)

Table 10: Arm Demographics, cont'd

		46 mg TPM (N=108)	92 mg TPM (N=107)	Qnexa 3.75/23 mg (N=241)
Sex	Female	86 (79.6)	85 (79.4)	201 (83.4)
	Male	22 (20.4)	22 (20.6)	40 (16.6)
Race	American Indian or Alaska Native	0 (0)	1 (0.9)	2 (0.8)
	Asian	1 (0.9)	1 (0.9)	2 (0.8)
	Black or African-American	11 (10.2)	22 (20.6)	39 (16.2)
	Native Hawaiian or other Pacific Islander	0 (0)	0 (0)	1 (0.4)
	White	93 (86.1)	82 (76.6)	192 (79.7)
	Multiple	2 (1.9)	1 (0.9)	0 (0)
	Other	1 (0.9)	0 (0)	5 (2.1)
Ethnicity	Hispanic or Latino	11 (10.2)	7 (6.5)	29 (12.0)
	Not Hispanic/Latino	97 (89.8)	100 (93.5)	212 (88.0)
Age	Mean (Std. Dev.)	46.4 (12.6)	45.3 (11.2)	42.5 (11.0)
Base Weight	Mean (Std. Dev.)	104.5 (15.6)	118.5 (21.9)	

Table 11: Arm Demographics, cont'd

		Qnexa 7.5/46 mg (N=605)	Qnexa 15/92 mg (N=1615)	All Arms (N=4510)
Sex	Female	434 (71.7)	1202 (74.4)	3386 (75.1)
	Male	171 (28.3)	413 (22.6)	1124 (24.9)
Race	American Indian or Alaska Native	0 (0)	10 (0.6)	26 (0.6)
	Asian	1 (0.9)	12 (0.7)	29 (0.6)
	Black or African-American	26 (23.9)	225 (13.9)	643 (14.3)
	Native Hawaiian or other Pacific Islander	0 (0)	4 (0.3)	9 (0.2)
	White	46 (72.5)	1336 (82.7)	3727 (82.6)
	Multiple	2 (1.8)	12 (0.7)	32 (0.7)
	Other	1 (0.9)	16 (1.0)	44 (1.0)
Ethnicity	Hispanic or Latino	9 (8.3)	219 (13.6)	215 (13.3)
	Not Hispanic/Latino	100 (91.7)	1396 (86.4)	1402 (86.7)
Age	Mean (Std. Dev.)	45.9 (11.6)	47.2 (12.1)	47.3 (11.7)
Base Weight	Mean (Std. Dev.)	102.5 (17.9)	106.6 (19.4)	106.4 (19.5)

A.1.2 Comparison of Qnexa 3.75/23 mg to Placebo

Table 12: OR for Subjects Experiencing Post-Baseline Event: Qnexa 3.75/23 mg against Placebo

	Qnexa 3.75/23 (N=241) N (%)	Placebo (N=1617) N (%)	Odds Ratio (95% CI)
Wish to be Dead	1 (0.41)	9 (0.56)	0.74 (0.09, 5.90)
Non-Specific Active Suicidal Thoughts	1 (0.41)	5 (0.31)	1.34 (0.16, 11.55)
Active Ideation: Methods, No Intent	0 (0)	2 (0.12)	*
Active Ideation: Intent, No Plan	0 (0)	0 (0)	*
Active Ideation: Plan & Intent	0 (0)	0 (0)	*
Composite Total (by Unique Subjects)	1 (0.41)	11 (0.68)	0.61 (0.08, 4.73)
Composite Total with Behaviors	1 (0.41)	12 (0.74)	0.56 (0.07, 4.31)

* Either arm having 0 subjects with events means OR cannot be calculated

Source: Calculated by reviewer

A.1.3 Post-Baseline Ideation Severity for Most Severe Ideation

These Tables (13 through 18) were constructed to contain only one record per subject; thus, any subject with more than one ideation was taken as the highest of the recorded severity or severity subscale ratings. The results from subjects responses for their "most common" suicidal ideation were not significantly different from those provided in the Tables. Note that in one or two subjects' cases, the severity subscales were not administered, for unknown reasons, resulting in a discepancy in total numbers from Table 3.

Table 13: Post-Baseline Ideation Severity: Subject's Most Severe Ideation, N (% of column total)

	Placebo	Qnexa 3.75/23 mg	Qnexa 7.5/46 mg	Qnexa 15/92 mg	All Qnexa
Wish to be Dead	6 (60.0)	1 (100.0)	2 (100.0)	13 (86.7)	16 (88.8)
Non-Specific Active Suicidal Thoughts	1 (10.0)	0 (0)	0 (0)	1 (6.7)	1 (5.6)
Active Ideation: Methods, No Intent	2 (20.0)	0 (0)	0 (0)	0 (0)	0 (0)
Active Ideation: Intent, No Plan	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Active Ideation: Plan & Intent	1 (10.0)	0 (0)	0 (0)	1 (6.7)	1 (5.6)
Total	10	1	2	15	18

Source: Calculated by reviewer.

A.1.4 Post-Baseline Controllability for Most Severe Ideation

Table 14: Most Severe Post-Baseline Ideation Controllability: N (% of column total)

	Placebo	Qnexa 3.75/23 mg	Qnexa 7.5/46 mg	Qnexa 15/92 mg	All Qnexa
Does not Attempt to Control Thoughts	0 (0)	0 (0)	0 (0)	1 (6.7)	1 (5.6)
Easily Able to Control Thoughts	10 (100.0)	0 (0)	2 (100.0)	10 (66.7)	12 (66.7)
Can Control with Little Difficulty	0 (0)	1 (100.0)	0 (0)	2 (13.3)	3 (16.7)
Can Control with Some Difficulty	0 (0)	0 (0)	0 (0)	1 (6.7)	1 (5.6)
Can Control with a Lot of Difficulty	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Unable to Control	0 (0)	0 (0)	0 (0)	1 (6.7)	1 (5.6)
Total	10	1	2	15	18

Source: Calculated by reviewer.

A.1.5 Post-Baseline Deterrents for Most Severe Ideation

Table 15: Most Severe Post-Baseline Ideation Deterrents: N (% of column total)

	Placebo	Qnexa 3.75/23 mg	Qnexa 7.5/46 mg	Qnexa 15/92 mg	All Qnexa
Does Not Apply	0 (0)	0 (0)	2 (66.7)	5 (33.3)	7 (36.8)
Deterrents Definitely Stopped From Attempting	10 (100.0)	1 (100.0)	1 (33.3)	10 (66.7)	12 (63.2)
Deterrents Probably Stopped From Attempting	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Uncertain Deterrents Stopped From Attempting	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Deterrents Most Likely Didn't Stop Attempt	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Deterrents Definitely Didn't Stop Attempt	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Total	10	1	2	15	18

Source: Calculated by reviewer.

A.1.6 Post-Baseline Duration for Most Severe Ideation

Table 16: Most Severe Post-Baseline Ideation Duration: N (% of column total)

	Placebo	Qnexa 3.75/23 mg	Qnexa 7.5/46 mg	Qnexa 15/92 mg	All Qnexa
Fleeting: Few Seconds or Minutes	8 (80.0)	1 (100.0)	2 (100.0)	8 (53.3)	11 (61.1)
Less Than an Hour/ Some of the Time	1 (10.0)	0 (0)	0 (0)	3 (20.0)	3 (16.7)
1–4 Hours/ A Lot of the Time	1 (10.0)	0 (0)	0 (0)	2 (13.3)	2 (11.1)
4–8 Hours/ Most of the Day	0 (0)	0 (0)	0 (0)	2 (13.3)	2 (11.1)
More Than 8 Hours/ Consistent/Continuous	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Total	10	1	2	15	18

Source: Calculated by reviewer.

A.1.7 Post-Baseline Frequency for Most Severe Ideation

Table 17: Most Severe Post-Baseline Ideation Frequency: N (% of arm)

	Placebo	Qnexa 3.75/23 mg	Qnexa 7.5/46 mg	Qnexa 15/92 mg	All Qnexa
Less Than Once per Week	9 (60.0)	0 (0)	2 (100.0)	8 (53.3)	10 (55.6)
Once/Week	1 (10.0)	0 (0)	0 (0)	2 (13.3)	2 (11.1)
2–5 Times/Week	0 (0)	1 (100.0)	0 (0)	5 (33.3)	6 (33.3)
Daily or Almost Daily	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Many Times/Day	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Total	10	1	2	15	18

Source: Calculated by reviewer.

A.1.8 Post-Baseline Reasons for Most Severe Ideation

Table 18: Most Severe Post-Baseline Ideation Reasons: N (% of arm)

	Placebo	Qnexa 3.75/23 mg	Qnexa 7.5/46 mg	Qnexa 15/92 mg	All Qnexa
Does Not Apply	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Completely to get Attention, Revenge, or a Reaction	2 (20.0)	0 (0)	0 (0)	2 (13.3)	2 (11.1)
Mostly to get Attention, Revenge, or a Reaction	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Equally for Attention, and to Stop/End Pain	0 (0)	1 (100.0)	0 (0)	0 (0)	1 (5.6)
Mostly to Stop/ End the Pain	5 (50.0)	0 (0)	1 (50.0)	10 (66.7)	11 (61.1)
Completely to Stop/ End the Pain	3 (30.0)	0 (0)	1 (50.0)	3 (20.0)	4 (22.2)
Total	10	1	2	15	18

Source: Calculated by reviewer.

A.1.9 Type of Worsening Suicidality by Treatment Arm

All categories are in comparison to a subject's baseline measurements. For example, a subject was considered to have an "Emerging Suicidal Ideation" if no ideation was recorded at baseline, but a suicidal ideation of any type was recorded during the periods when the subject was on treatment (or within 28 days of last treatment).

Table 19: Type of Worsening Suicidality by Arm, N (% of Arm)

	Placebo (N=1617)	Qnexa 3.75/23 mg (N=241)	Qnexa 7.5/46 mg (N=605)	Qnexa 15/92 mg (N=1615)	All Qnexa (N=2461)
Emergence of Ideation	7 (0.43)	1 (0.41)	2 (0.33)	10 (0.62)	13 (0.53)
Emergence of Behavior	2 (0.12)	0 (0)	0 (0)	1 (0.06)	1 (0.04)
Higher Level Ideation	0 (0)	0 (0)	0 (0)	1 (0.06)	1 (0.04)
Increase in at Least One Severity Marker	2 (0.12)	0 (0)	0 (0)	0 (0)	0 (0)

Source: Calculated by reviewer.

Signatures/Distribution List

Primary Statistical Reviewer: Benjamin Neustifter, Ph.D.
Date: May 17, 2010

Division Director: Aloka Chakravarty, Ph.D.

Acting Statistical Team Leader: Mat Soukup, Ph.D.

cc:

Archival NDA
HFD-700/Chakravarty
HFD-700/Levenson
HFD-700/Neustifter
HFD-700/Soukup
HFD-510/Colman
HFD-510/Dharia
HFD-510/Egan
HFD-510/Madara
HFD-510/Parks
HFD-510/Roberts
HFD-700/Patrician

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-22580	ORIG-1	VIVUS INC	QNEXA (phentermine IR + topiramate modified release) CAPSULE; VI-0521

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

/s/

BENJAMIN B NEUSTIFTER
05/17/2010
LaTeX mis-built the file last time and somehow included the old signature page. It's fixed now.

ALOKA G CHAKRAVARTY
05/17/2010

MATTHEW J SOUKUP
05/17/2010
Concur with review