CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

200063Orig1s000

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

SECONDARY STATISTICAL REVIEW MEMO

CLINICAL STUDIES

NDA/Serial Number:	200063 /No. 42
Drug Name:	Contrave™ (naltrexone / bupropion) extended-release tablets
Indication(s):	Weight management
Applicant:	Orexigen
Date(s):	Submission date: December 11, 2013
Review Priority:	Standard
Biometrics Division:	Division of Biometrics 2 (HFD-715)
Secondary Statistical Reviewer:	Mark Rothmann, Ph.D. , Statistical Team Leader
Concurring Reviewer:	Thomas Permutt, Ph.D., Division Director, Division of Biometrics 2
Medical Division:	Division of Metabolism and Endocrinology Products (HFD-510)
Clinical Team:	Eileen Craig, M.D., Medical Reviewer Jim Smith, M.D., Medical Team Leader Eric Colman, M.D., Medical Deputy Division Director Jean-Marc Guettier, M.D., Division Director
Project Manager:	Pat Madara

Keywords: Missing Data, Obesity Study

1. EXECUTIVE SUMMARY

This resubmission of NDA 200063 dated December 10, 2013 is in response to the January 31, 2011 Complete Response Letter. Reference is made to the statistical reviews of Dr. Janice Derr signed December 15, 2010 and Dr. Lee-Ping Pian signed May 14, 2014.

It is important to provide patients and prescribers the best quality of information in the product label on the effects on a treatment, so that informed treatment decisions can be made. This review provides a recommendation on the analysis that best describes the treatment effect and that should appear in the product label.

1.1 Conclusions and Recommendations

We are recommending that the sponsor's ITT analysis be the analysis provided in the label for providing the treatment effect. The treatment effect is not appropriately represented by the FAS analysis. That analysis excludes subjects who could not tolerate 4 weeks of Contrave (the first four weeks is a titration period). Those subjects had very little change in weight from baseline to 56 weeks and the primary analysis inappropriately has their weight change represented by subjects who could tolerate 4 weeks of Contrave. The FAS analysis has the weight change of subjects who could not tolerate the study drug represented by the weight change of subjects who could tolerate the study drug. The hypothetical benefit to subjects who could not tolerate study drug, if only they could tolerate it, does not a meaningfully represent the effect of the study drug.

The sponsor's ITT analysis is also similar to the analysis procedure that is in the Qsymia label.

1.2 Background and Results

Section 14 (Clinical Studies) of the applicant's proposed label includes results from studies NB-301, NB-302 and NB-304. These were randomized, placebo-controlled studies that compared Contrave to placebo. There were co-primary endpoints of (1) body weight percent change from baseline to Week 56 and (2) the proportion of patients with at least 5% weight loss at Week 56. The primary analysis population was the Full Analysis Set (FAS) defined as all subjects who were randomized, had a baseline body weight measurement, and at least one post-baseline body weight measurement while on study drug. The primary analysis only considered body weight measurements while on study drug and utilized a last observation carried forward (LOCF) imputation of missing week 56 body weight measurements. A secondary analysis was based on an analysis set (referred by the applicant as the intent-to-treat analysis set) which included all randomized subjects with a baseline and post baseline body weight, where the endpoint was defined as the last non-missing post baseline measurement during the double-blind treatment phase (irrespective of being on study drug at the time of the last measurement). This secondary analysis is similar to the primary analysis for the Qsymia studies (see below).

The results from these two analyses are provided in the table below.

Analysis Results - Changes in Weight in 56-Week Trials

	COR-I		COR-BMOD		COR-Diabetes	
	CONTRAVE 32 mg/360 mg	**Placebo**	**CONTRAVE 32 mg/360 mg**	**Placebo**	**CONTRAVE 32 mg/360 mg**	**Placebo**
FAS N**	471	511	482	193	265	159
Baseline mean (SD)	100.2 (16.3)	99.3 (14.3)	100.7 (15.4)	101.9 (15.0)	106.4 (19.1)	105.0 (17.1)
LS Mean % Change From Baseline (SE)	-6.1 (0.3)	-1.3 (0.3)	-9.3 (0.4)	-5.1 (0.6)	-5.0 (0.3)	-1.8 (0.4)
Difference from placebo (95% CI)	-4.8 (-5.6, -4.0)		-4.2 (-5.6, -2.9)		-3.3 (-4.3, -2.2)	
ITT* N	538	536	565	196	321	166
Weight (kg)						
Baseline mean (SD)	99.8 (16.1)	99.5 (14.4)	100.3 (15.5)	101.8 (15.0)	104.2 (19.1)	105.3 (16.9)
LS Mean % Change From Baseline (SE)	-5.4 (0.3)	-1.3 (0.3)	-8.1 (0.4)	-4.9 (0.6)	-3.7 (0.3)	-1.7 (0.4)
Difference from placebo (95% CI)	-4.1 (-4.9, -3.3)		-3.2 (-4.5, -1.8)		-2.0 (-3.0, -1.0)	
All Randomized Subjects, N	583	581	591	202	335	170

Notice that the FAS-based analysis excludes 12%-17% of subjects on the Contrave arm in the ITT population, and excludes 1.5%-5% of subjects on the placebo arm in the ITT population. Additionally, notice that the FAS-based analysis excludes 18%-21% of subjects randomized to the Contrave arm, and excludes 4%-12% of subjects randomized to the placebo arm. The clinical trials of Contrave consistently had a much larger fraction of subjects who did not have a post baseline measurement while on treatment in the Contrave group than in the placebo group.

Part of a therapy's effect is mediated through the ability to tolerate the therapy. Also, on average subjects who discontinued Contrave and later had their weight measured at 56 weeks had little weight change from baseline to 56 weeks. Therefore, an analysis that excludes subjects who could not tolerate the therapy and weight measurement after stopping therapy would likely produce a higher estimate than this effect. When the known weight measurements are taken into consideration, as in the applicant's secondary analysis, the difference in mean percent weight change is consistently less than that estimated by the FAS-LOCF across studies.

Again, I believe that missing data should be addressed in the analysis in the most appropriate way to provide the most relevant estimate of the treatment difference/effect. I believe that the applicant's secondary analysis (i.e, their analysis based on the "intent-to-treat analysis set") should be provided in the product label, (b) (4)

The Agency's thinking on how to address missing has also evolved since the publication of the 2010 report on missing data by the National Academy of Sciences, *The Prevention and Treatment of Missing Data in Clinical Trials*. The FDA commissioned this report.

See Dr. Derr's review for further information on the design of these studies and a summary of the results.

1.3 Analyses used for other products

There was interest about the analyses that were performed in the reviews of Qsymia and Belviq and that appear in the product labels for those products. Below is a summary for Qsymia and Belviq.

Qsymia (Approved 7/17/2012; Statistical review signed 9/27/2010)

Reference is made to the statistical review of Dr. Lee-Ping Pian signed 9/27/2010 and to the Qsymia product label.

- Co-Primary Endpoints: (1) Body weight percent change from baseline to Week 56 and (2) the proportion of patients with at least 5% weight loss at Week 56.

- Per the statistical review for each study: "The sponsor's primary efficacy analysis used the ITT population which included both on-drug and off-drug patients."

- The analyses included in the product label: "uses all available data from subjects in ITT population, including data collected from subjects who discontinued drug but remained on study. LOCF method used to impute missing data." (Separate analyses for each study)

- The only analysis on body weight percent change from baseline in the statistical review uses the MITT population and only included data while patients were on study drug.

Belviq (Approved 6/27/2012; Statistical review signed 9/22/2010)

Reference is made to the statistical review of Dr. Janice Derr signed 9/22/2010 and to the Belviq product label.

- Primary endpoints in hierarchical order: (1) proportion of subjects achieving at least 5% reduction in body weight at the end of year one, (2) body weight percent change from baseline to Week 52, and (3) the proportion of subjects achieving at least 10% reduction body weight at the end of year one.

- Per the statistical review for each study (Study 009 and Study 011), the primary analysis on body weight percent change from baseline to Week 52 was based on the MITT

population, included only measurements while on drug, and uses LOCF to impute missing data.

- The primary analysis results for Study 009 and Study 011 are not (separately) in the product label. Instead, a combined analysis is included in the product label for all of the primary endpoints. For the combined analysis on body weight percent change from baseline to Week 52 the label states:
    - "Intent to Treat Population using last observation carried forward method; All patients who received study medication and had a post-baseline body weight." And
    - that the analysis method is "Least squares means adjusted for baseline value, treatment, study and treatment by study interaction."
    - Comments on the analysis
        - This is not an analysis that combines the results of the individual primary analyses of the studies.
        - There is no baseline value by study interaction term. This means the same prognostic value for baseline weight is forced on both studies. This is different from using baseline weight as covariate in the individual study analyses, where the prognostic value of baseline weight is allowed to be different between the studies.
        - Having a treatment by study interaction term means that there will be different estimated treatment effects for the two separate studies. The label does not provide any information on how these study-specific estimated treatment effects are combined to get the one estimate provided in the labeling.

- There were many additional analyses in the statistical review. These include analyses using measurements while off study drug (with LOCF) and analyses that excluded over 50% of the subjects. The statistical review does not contain any combined analyses across studies.

For every study across all three products when measurements while off study drug are included in the analysis, the "treatment difference" (which is represented as a negative value) moves towards zero.

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

/s/

MARK D ROTHMANN
09/08/2014

THOMAS J PERMUTT
09/08/2014
I concur.

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

STATISTICAL MEMORANDUM

CLINICAL STUDIES

NDA/Serial Number: 200063 /No. 42, 45

Drug Name: Contrave™ (naltrexone / bupropion) extended-release tablets

Indication(s): Weight management

Applicant: Orexigen

Date(s): Submission date: December 11, 2013, February 7, 2014
PDUFA Goal Date: June 11, 2014

Review Priority: Standard

Biometrics Division: Division of Biometrics 2 (HFD-715)

Statistical Reviewer: Lee Ping Pian, Ph.D.

Concurring Reviewers: Mark Rothmann, Ph.D. , Statistical Team Leader
Tom Permutt, Ph.D., Division Director

Medical Division: Division of Metabolism and Endocrinology Products (HFD-510)

Clinical Team: Eileen Craig, M.D., Medical Reviewer
Jim Smith, M.D., Medical Team Leader
Eric Colman, M.D., Medical Deputy Division Director
Jean-Marc Guettier, M.D., Division Director

Project Manager: Pat Madara

Keywords: labeling review, clinical study

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

For labeling purposes, we examined the weight loss from the phase 3 clinical studies and the body weight data from interim analysis of study NB-CVOT. The recommendation is that patients should stop taking NB if weight loss is less than 3% by week 16 of NB treatment.

Brief Overview

The resubmission dated December 10, 2013 was a complete response to January 31, 2011 Complete Response Letter. Section 3.1[1] of the December submission, Treatment Algorithm (Week 16 Assessment) Justification, of the Safety Update stated that "The original NDA includes a proposed treatment algorithm to provide direction to prescribers regarding continuation of NB32 treatment in the event that a patient does not lose a clinically meaningful amount of body weight at a specified time point after initiating treatment… the most appropriate threshold to support continued long-term treatment with NB32 was achieving at least 5% weight loss by Week 16." Section 3.1.1, Analysis of Phase 3 Body Weight Data, presented analyses based on Agency recommendations (the CRL and at the March 11, 2013 Type C meeting[2]). The sponsor concluded that "These results, suggest that the timeframe for evaluation (16 weeks) and the 5% criterion for clinically important early weight change could provide a reasonable combination of early time point and weight loss threshold to inform the appropriateness of continued treatment."

The sponsor, however also stated that "data from Study NB-CVOT that included a prospectively defined weight loss criterion for continuation of treatment supports further evaluation and possible modification of the treatment algorithm." The safety update[3] of the February 7, 2014 submission included Justification for Removal of Proposed Treatment Algorithm. The sponsor stated that "The NDA includes a proposed treatment algorithm to provide direction to prescribers regarding continuation of CONTRAVE (NB32) treatment. This was proposed by Orexigen in large part as a safety measure to avoid inappropriate exposure to NB32 for those subjects who may not exhibit the expected degree of weight loss and to balance any theoretical cardiovascular (CV) risk related to the small increases in blood pressure that accompany NB32 treatment in some patients. Specifically, the retrospective Phase 3 analyses, which were targeted to optimally identify patients losing ≥5% bodyweight at one year, suggested that the most appropriate threshold to support continued long-term

[1] Link in appendix

[2] See appendix

[3] Link in appendix

treatment with NB32 was achieving at least 5% weight loss by Week 16. However, and as noted in the NDA Resubmission Safety Update, the treatment algorithm assumptions have been re-evaluated in light of the results of Study NB-CVOT…" and "Importantly, the interim analysis intended to support approval of NB32 suggests no unique risk to subjects who continue long-term treatment even when a weight loss threshold substantially lower than the previously proposed 5% is applied…"

The February submission contains the interim analysis of the cardiovascular outcome study NB-CVOT. Using the one year weight loss data, I performed similar sensitivity and specificity analysis as requested by FDA to identify a weight loss threshold at early week to predict the likelihood of weight loss <5% at week 52 for both Belviq and Qsymia labelings. The method has its limitations due to the proposed plans for discontinuing therapy was not prospectively studied. There was no randomization to either continuing therapy or discontinuing therapy.

Study Results

Study NB-CVOT designed prospectively withdraw from study medication those subjects achieving less than 2% weight loss at Week 16 or experiencing consecutive, sustained increases in BP (SBP or DBP) of ≥ 10 mmHg. Table 1 displays sensitivity and specificity by treatment group (observed cases or LOCF) using week 16 weight loss threshold (2%, 3%, 4% and 5%) to predict week 52 weight loss of <5% or ≥ 5%. As defined by FDA, the numerator of sensitivity is the number of patients who lost less than the threshold at week 16 and who lost less than 5% at week 52. The denominator of sensitivity is the number of patients whose weight loss was less than the threshold. The numerator of specificity is the number of patients whose weight loss was greater than or equal to the threshold and greater than or equal to 5% at week 52. The denominator of specificity is the number of patients whose weight loss was greater than or equal to the threshold at week 16. Table 1 showed that the sensitivity is high (88%, OC) for the 3% threshold.

The percentage of patients who had ≥ 5% body weight loss at week 52 was 43% (863/2020) for NB and 23% (448/1925) for Placebo (OC). For LOCF, it was 33% (1433/4376) for NB and 13% (563/4370) for placebo.

Table 1 Sensitivity and specificity for 5% weight loss by week 16 wt loss threshold – Treatment NB, Study CVOT

	LOCF NB n=4376		LOCF Placebo n=4370	
x% wt loss threshold At Week 16	Sensitivity*	Specificity**	Sensitivity*	Specificity**
2%	1627/1629 (100%)	1431/2747 (52%)	2648/2655 (100%)	556/1715 (32%)
3%	2152/2207 (98%)	1378/2169 (64%)	3224/3308 (97%)	479/1062 (45%)
4%	2480/2627 (94%)	1286/1749 (74%)	3511/3653 (96%)	421/717 (59%)
5%	2727/3004 (91%)	1156/1372 (84%)	3663/3891 (94%)	335/479 (70%)

	Observed cases NB n=2020		OC Placebo n=1925	
x% wt loss threshold At week 16	Sensitivity*	Specificity**	Sensitivity*	Specificity**
2%	468/508 (92%)	823/1512 (54%)	881/988 (89%)	341/937 (36%)
3%	674/770 (88%)	767/1250 (61%)	1142/1315 (87%)	275/610 (45%)
4%	821/990 (83%)	694/1030 (67%)	1280/1494 (86%)	234/431 (54%)
5%	924/1193 (77%)	594/827 (72%)	1360/1630 (83%)	178/295 (60%)

*Sensitivity: $\frac{\text{(\# with both wt loss <5\% at week 52 and wt loss < x\% at week 16)}}{\text{(\# wt loss<x\% at week 16)}}$

**Specificity: $\frac{\text{(\# with both wt loss ≥5\% at week 52 and wt loss ≥ x\% at week 16)}}{\text{(\# wt loss ≥ x\% at week 16)}}$

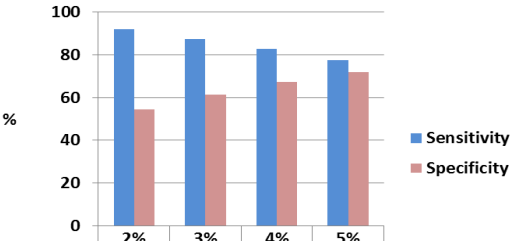
Table 2 displays proportion of patients with weight loss ≥5% at week 52 given they lost x% to y% of weight at week 16 (column 1).

Table 2 Proportion of weight loss ≥5% at Week 52 for given Weight loss at Week 16 (OC)

x% to y% wt loss At week 16	weight loss ≥5% at week 52	
	NB	Placebo
2% - 3%	56/262 (21%)	66/327 (20%)
3% - 4%	73/220 (33%)	41/179 (23%)
4% - 5%	100/203 (49%)	56/136 (41%)

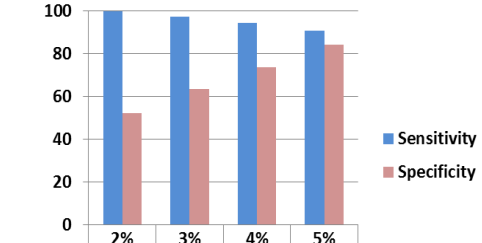
Treatment NB

OC



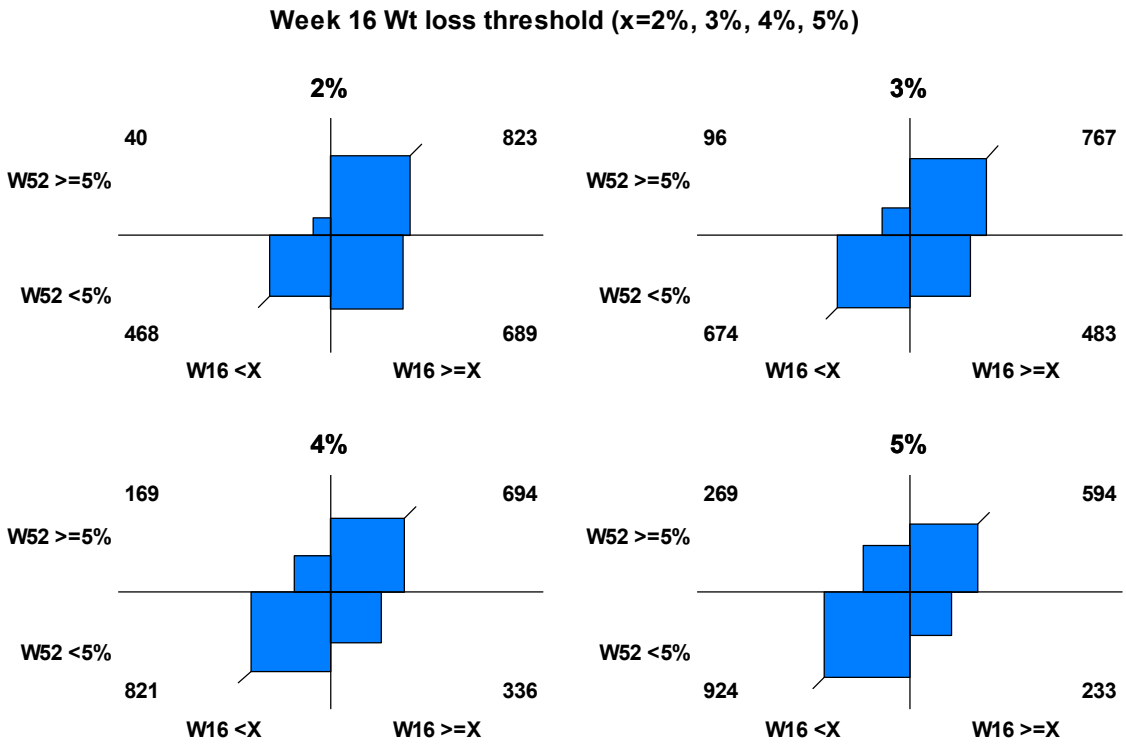
	2%	3%	4%	5%
Sensitivity	92	88	83	77
Specificity	54	61	67	72

LOCF

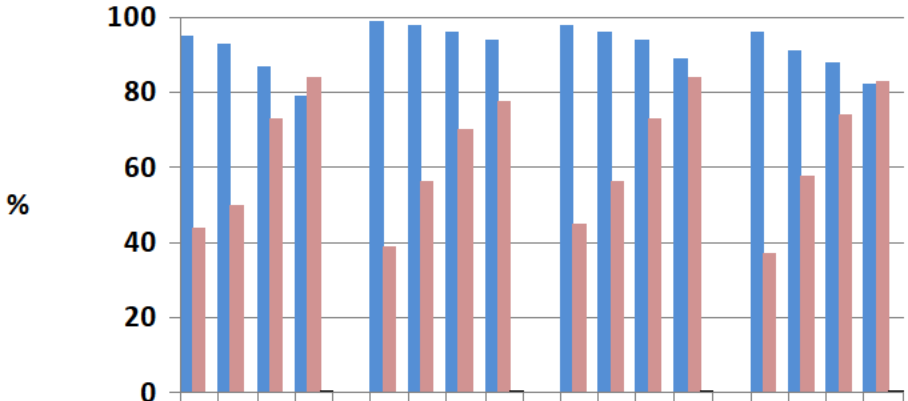


	2%	3%	4%	5%
Sensitivity	100	98	94	91
Specificity	52	64	74	84

The graph below depicts two-by-two table of body weight loss strata for Week 16 (2%, 3%, 4% and 5%) vs. week 52 weight loss strata of 5%. The number of patients at the lower left hand corner when divided by the total number of the left hand side column corresponds to sensitivity. The number of patients at the upper right hand corner when divided by the total number on the right hand side column corresponds to specificity. The upper left hand corner represents patients lost less than threshold at week 16 but lost ≥ 5% weight at week 52.

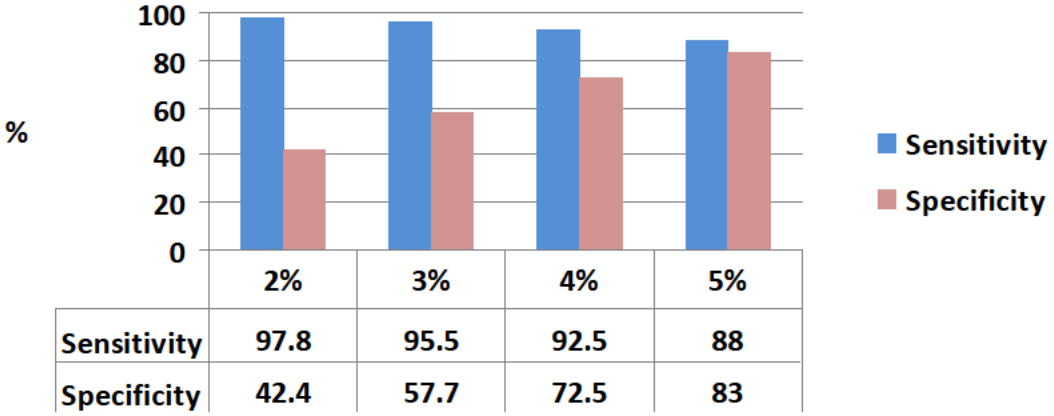


Phase 3 studies: The sample sizes (LOCF) for S1 to S4 were 471, 482, 825 and 265, respectively.



	S1				S2				S3				S4			
Sensitivity	95	93	87	79	99	98	96	94	98	96	94	89	96	91	88	82
Specificity	44	50	73	84	39	56	70	78	45	56	73	84	37	58	74	83
Threshold	2%	3%	4%	5%	2%	3%	4%	5%	2%	3%	4%	5%	2%	3%	4%	5%

Phase 3 pooled study



	2%	3%	4%	5%
Sensitivity	97.8	95.5	92.5	88
Specificity	42.4	57.7	72.5	83

Conclusions and Recommendations

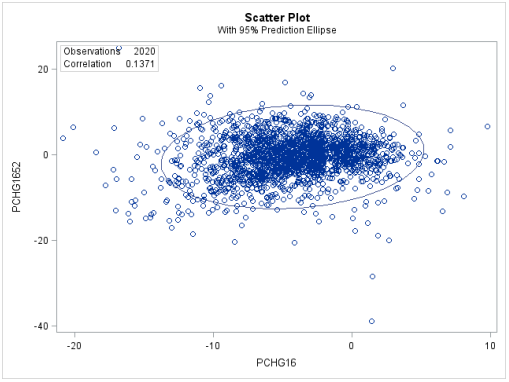
The body weight loss of <3% at week 16 is recommended to identify patients who are less likely to have ≥ 5% weight loss at week 52.

Additional analyses – Study NB-CVOT

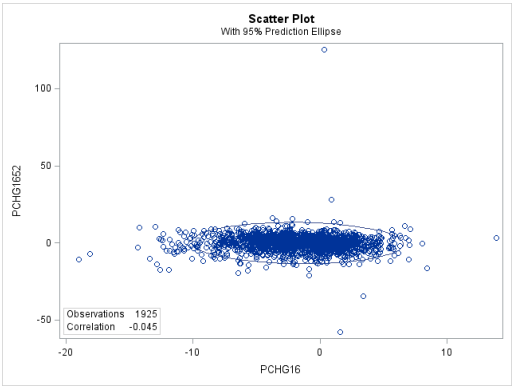
Figure 2 displays the percent weight change difference (Week 52 – Week 16) vs. percent weight change at week 16 for the observed cases. The correlation was +0.14 (R^2 = 0.02) for treatment NB and -0.05 for placebo. Only 2% of the variation for percent change from week 16 to week 52 can be explained by the linear model for the NB treatment group. For each treatment arm, there was little correlation between the percent weight change at week 16 and any additional percent weight change by week 52.

Figure 1 % weight change from baseline Scatter plots – (Week 52-Week 16) vs. Week 16

Treatment=NB



Placebo



The mean percent weight change from week 16 to week 52 was -0.43% for NB and -0.04% for placebo. See Table 2 for further details.

Table 2 Percent weight change at week 16 and percent weight change from week 16 to week 52

Estimate	NB n=2020	Placebo n=1925
Mean (SD) % weight change at week 16 [min, max]	-4.3 (3.9) [-20.8, 9.7]	-1.8 (3.4) -19.0, +13.9
% weight change from week 16 to week 52 [min, max]	-0.43 (5.0) [-38.9, 25.1]	-0.045 (5.5) [-57.8, +125.3]
Correlation	+0.14	-0.05

Appendix

Statistical request (March 13, 2013 Type C meeting)

"Tabulate the sensitivity and specificity for identifying 5% non-responders at the week 56 endpoint for each of the four Phase 3 studies separately, using the same range of early visits and early weight loss threshold that are shown in Table 2, Appendix 5 of the briefing document. We request that sensitivity and specificity be re-defined so that sensitivity refers to correctly identifying a 5% non- responder at week 56, based on being classified as an "early weight loss [*threshold level]*" non- responder. Continue to use the FAS/LOCF analysis sets as you did in the briefing document."

Safety update link (December 11, 2013 submission)
(\\cdsesub1\evsprod\nda200063\0041\m5\53-clin-stud-rep\535-rep-effic-safety-stud\obesity\5353-rep-analys-data-more-one-stud\safety-update\safety-update.pdf)

Safety update link (February 7, 2014 submission)
\\cdsesub1\evsprod\nda200063\0044\m5\53-clin-stud-rep\535-rep-effic-safety-stud\obesity\5353-rep-analys-data-more-one-stud\safety-update\safety-update.pdf

---

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
05/14/2014

MARK D ROTHMANN
05/14/2014
I concur

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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA:	200063 (Resubmission)
Supplement #:	SDN 42 (eCTD Sequence No. 0041)
Drug Name:	Naltrexone SR 32 mg/Bupropion SR 360 mg tablets
Indication(s):	Treatment of obesity and weight management
Applicant:	Orexigen Therapeutics, Inc.
Date(s):	Receipt Date: December 11, 2013
	PDUFA Date: June 11, 2014
Review Timeline:	6 months
Biometrics Division:	Division of Biometrics VII (DBVII)
Statistical Reviewer:	Janelle K. Charles, PhD
Concurring Reviewers:	Mat Soukup, PhD
	Team Leader, DBVII
	Aloka Chakravarty, PhD
	Director, DBVII
Medical Division:	Division of Metabolism and Endocrinology Products (DMEP)
Clinical Team:	Eileen Craig, MD
	Medical Officer, DMEP
	James Smith, MD
	Medical Team Leader, DMEP
Project Manager:	Patricia Madara
	Regulatory Project Manager, DMEP

Keywords: CV safety, trial integrity/blinding, interim analysis, Cox model

Table of Contents

LIST OF ABBREVIATIONS **4**

LIST OF TABLES **5**

LIST OF FIGURES **6**

1 EXECUTIVE SUMMARY **7**

1.1 INTERIM ANALYSIS FINDINGS AND PRE-APPROVAL CONCLUSIONS 8
1.2 STATISTICAL ISSUES AND POST-APPROVAL IMPLICATIONS 9
1.2.1 High Percentage of Treatment Discontinuations 9
1.2.2 Unblinding and LIGHT Trial Integrity 9

2 INTRODUCTION **11**

2.1 OVERVIEW AND REGULATORY BACKGROUND 11
2.2 DATA SOURCES 14

3 STATISTICAL SAFETY EVALUATION **14**

3.1 DATA AND ANALYSIS QUALITY 14
3.1.1 Pre-Approval Findings: Quality of Interim Data and Interim Study Report 14
3.1.2 Post-Approval Implications: Unblinding and Trial Integrity 15
3.2 STUDY DESIGN AND ENDPOINTS 16
3.2.1 Study Design 16
3.2.2 Definition of Endpoints and Adjudication Process 19
3.3 STATISTICAL METHODOLOGIES 20
3.3.1 Analysis Population 21
3.3.2 Statistical Hypotheses and Sample Size Estimation 21
3.3.3 Applicant's Planned Statistical Analyses 22
3.3.4 Reviewer's Post Hoc Analyses 22
3.4 SUBJECT DISPOSITION, DEMOGRAPHIC AND BASELINE CHARACTERISTICS 23
3.5 RESULTS OF STATISTICAL ANALYSES 28
3.5.1 Results of MACE Analyses 28
3.5.2 Results of MACE+ Analysis 30
3.5.3 Results of MACE Component Analyses 31
3.5.4 All-cause Mortality Analysis Results 31

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS **31**

4.1 GENDER, RACE, AND AGE 31
4.2 OTHER SPECIAL/SUBGROUP POPULATIONS 32

5 SUMMARY AND CONCLUSIONS **34**

5.1 COLLECTIVE EVIDENCE AND STATISTICAL ISSUES 34
5.1.1 Collective Evidence from Pre- Approval Interim Analysis 34
5.1.2 Statistical Issues and Post-Approval Implications 35
5.2 CONCLUSIONS AND RECOMMENDATIONS 36

5.3 LABELING RECOMMENDATIONS 37

APPENDIX 1 ASSESSMENT OF PROPORTIONAL HAZARDS ASSUMPTION 38

APPENDIX 2 ADDITIONAL SUBJECTS INCLUDED IN INFORMATIVE CENSORING SENSITIVITY ANALYSES 39

APPENDIX 3 ADDITIONAL SUBJECTS INCLUDED IN ALL-CAUSE MORTALITY ANALYSIS 41

APPENDIX 4 PROJECTIONS FOR ON-TREATMENT EVENTS AT TRIAL COMPLETION 42

APPENDIX 5 LIST OF INDIVIDUALS WITH ACCESS TO UNBLINDED INTERIM DATA 43

SIGNATURE/DISTRIBTUION LIST 46

LIST OF ABBREVIATIONS

CDER	Center for Drug Evaluation and Research
CEC	Clinical Endpoint Committee
CI	Confidence Interval
CRL	Complete Response Letter
CVOT	Cardiovascular Outcome Trial
DAP	Data Access Plan
DMC	Data Monitoring Committee
EOR	End of Review
FDRR	Formal Dispute Resolution Request
HR	Hazard Ratio
ITT	Intent to Treat
MACE	CV Death, Myocardial Infarction, Stroke
MACE+	CV Death, Myocardial Infarction, Stroke, Unstable Angina
MI	Myocardial Infarction
ODEII	Office of Drug Evaluation II
OND	Office of New Drugs
PH	Proportional Hazards
PP	Per Protocol
SAP	Statistical Analysis Plan

LIST OF TABLES

Table 1 Summary of Findings from MACE Analyses..... 9

Table 2 Levels of Access to Unblinded Data for CVOT..... 16

Table 3 Distribution of Subject Trial Status by Treatment Group..... 24

Table 4 Summary of Reasons for Treatment Discontinuation..... 25

Table 5 Applicant's Projection of Cumulative Medication Retention Rates..... 26

Table 6 Distribution of Demographic Characteristics by Treatment (ITT Population)..... 27

Table 7 Distribution of Baseline CV Risk Factors by Treatment (ITT Population)..... 28

Table 8 Individual MACE Component Analyses Results..... 31

Table 9 Results of MACE Analyses by Gender, Race, and Age..... 32

Table 10 Results of MACE Analyses by CV Risk Factors..... 33

Table 11 Summary of Interim Analysis Results..... 35

LIST OF FIGURES

Figure 1 CVOT Design and Visit Schedule..... 18

Figure 2 CVOT Dosing Schedule..... 18

Figure 3 Incidence of MACE over Time by Treatment Group 29

Figure 4 Incidence of MACE+ over Time by Treatment Group 30

1 EXECUTIVE SUMMARY

This is a statistical safety review of a pre-approval interim analysis of an ongoing event-driven trial designed to assess the risk of major adverse cardiovascular (CV) events with CONTRAVE compared to placebo. CONTRAVE is a fixed combination of naltrexone hydrochloride and bupropion hydrochloride. Both components are currently approved monotherapies in the United States: naltrexone for the treatment of opiate and alcohol dependence, and bupropion for the treatment of major depression and nicotine dependence. The Applicant, Orexigen Therapeutics Inc., originally submitted the New Drug Application (NDA 200063) for CONTRAVE in March 2010, for the proposed indication of treatment of obesity and weight management, including weight loss, and maintenance of weight loss. FDA issued a complete response letter (CRL) on January 31, 2011 stating that before the application could be approved the Applicant "must conduct a randomized, double-blind, placebo-controlled trial of sufficient size and duration to demonstrate that the risk of major adverse cardiovascular events in overweight and obese subjects treated with naltrexone/bupropion does not adversely affect the drug's benefit-risk profile"[1].

Subsequently, the Applicant submitted three consecutive Formal Dispute Resolution Requests (FDRRs) to the Office of Drug Evaluation II (ODEII), the Office of New Drugs (OND), and the Center for Drug Evaluation and Research (CDER) that queried, among other issues, whether the requirement for a pre-approval CV outcomes trial was consistent with recent FDA actions on similar products. All of these FDRRs were denied[2,3,4]; however, the OND denial letter outlined key design and analysis parameters for the required trial. In particular, the letter recommended that the trial be designed such that a pre-approval hazard ratio (HR) risk margin of 2.0 can be ruled out, which would require at least 87 or approximately 25% of the planned events for the trial. The letter also stipulated that after approval, the trial demonstrates that a HR of 1.4 can be ruled out. These HR risk margins are based on the upper bounds of 95% (two-sided) confidence intervals (CIs). The interim analysis report and data for the required CV outcomes trial (CVOT) to assess the pre-approval HR risk margin of 2.0 that are contained in the NDA resubmission[5] is the subject of this review. Note that the CVOT is ongoing at the time of this statistical review; refer to Section 1.2 and 3.1 for detailed discussions of data quality and trial integrity issues. The pre-approval interim analysis comprises approximately 1.5 years of trial follow-up data.

The protocol for the required CVOT, also referred to as the LIGHT trial, was reviewed under Special Protocol Assessment[6] and followed the recommendations outlined in the OND denial letter. The LIGHT trial is multicenter, double-blind, 1:1 randomized, placebo-controlled, and includes overweight and obese subjects at increased risk of CV outcomes including subjects with

[1] Refer to Complete Response letter dated January 31, 2011.
[2] Refer to FDR Denial Letter by Dr. Curtis Rosebraugh, Director of ODE 2 dated July 7, 2011.
[3] Refer to FDR Denial Letter by Dr. John Jenkins, Director of OND dated September 15, 2011.
[4] Refer to FDR Denial Letter by Dr. Douglas Throckmorton of OND dated October 15, 2012.
[5] The NDA was resubmitted December 10, 2013 and the interim analysis report and analysis datasets were submitted on February 7, 2014, which is within 60 days of the resubmission date as stipulated in agreements with FDA.
[6] Refer to Special Protocol Agreement letter dated February 3, 2012.

type 2 diabetes (T2DM) and/or history of CV disease; refer to Section 3.2.1 for the specific definition of increased CV risk and more details about the trial design. The agreed upon primary analysis population is all randomized subjects, who are dispensed medication. The primary endpoint is **MACE,** a composite endpoint comprising CV death, non-fatal myocardial infarction (MI), or non-fatal stroke. The primary analysis is time from randomization to first MACE based on a Cox proportional hazard model; this analysis includes events that occurred after randomization through analysis cut-off date regardless if subjects discontinued treatment (referred throughout this review as the on-study analysis). On-treatment analysis[7] of MACE that contain only those events that occurred while subject is on treatment or within 30 days of treatment discontinuation is also presented in this review. All events included in the analyses are based on positively adjudicated events determined by an independent blinded committee that used standardized definitions. Because this review is based on the pre-approval interim analysis only and does not address any planned interim testing of the post-approval 1.4 HR margin, for which the trial is ultimately designed to test, all confidence intervals (CIs) presented in this review are based on 95% (two-sided) levels as no alpha adjustments are made for testing the 2.0 HR risk margin.

1.1 Interim Analysis Findings and Pre-Approval Conclusions

The primary analysis population consists of 4455 CONTRAVE subjects and 4450 placebo subjects. The distributions of demographic and CV risk factors are similar between the treatment groups. The majority of subjects are female (55%) or White (84%). The average age of subjects is 61 years and average BMI is 37 kg/m^2. Most subjects had T2DM (85%) and 32% of subjects had history of CV disease. Note that trial enrollment was complete at the time of the NDA resubmission; as such, the composition and baseline characteristics of the primary analysis population is not expected to change at the end of the trial.

Based upon the pre-approval interim analysis, the incidence of MACE during the trial is (b) (4) compared to placebo subjects (b) (4) The estimated HR for MACE is (b) (4) for the on-study analysis. The incidence of MACE while subjects are on-treatment or within 30 days of treatment discontinuation (b) (4) and the upper bound of the 95% CI is less than 2.0; see Table 1. Therefore, it can be concluded that the interim analysis results ruled out the pre-approval HR risk margin of 2.0. Refer to Section 3.5 for results of analyses for secondary and other endpoints assessed in this review.

It is important to note that this review is limited to interim data only which contains approximately a quarter of the total planned number of events for the trial. Therefore, these results may not reflect findings observed at the conclusion of the trial (b) (4)

[7] The sample size and power for the LIGHT trial is based on the on-study analysis. According to the End of Review meeting minutes dated June 27, 2011; consistent results of the on-study and on-treatment analyses are expected.

(b) (4)

Table 1 Summary of Findings from MACE Analyses

	CONTRAVE n/N (%)	Placebo n/N (%)	HR (95% CI)
Primary (On-study) analysis[1]	(b) (4)		
On-treatment analysis[2]			

[1] Includes all MACE that occurred while subject is followed in the trial, regardless of treatment status
[2] Includes only those MACE that occurred while subject is still being treated or within 30 days of treatment discontinuation.
n=number of subjects with MACE, N=number of randomized subjects, HR=hazard ratio from Cox PH model, CI=confidence interval

Source: Created by the statistical reviewer using dataset "adtte.xpt"

Regarding the CONTRAVE label, the recommendation is that none of the findings of the pre-approval interim analyses of the LIGHT trial be included in the label. Additionally, because the CV safety of CONTRAVE has not been confirmed with this interim data, the recommendation is that a limitation of use statement with respect to CV morbidity and mortality be included in the label. An example of such statement is "The effect of CONTRAVE on cardiovascular morbidity and mortality has not been established", which is consistent with labels for other currently approved weight loss products, namely, Belviq and Qsymia. If such a statement is included in the CONTRAVE label, the recommendation is that it be revisited after submission of the findings from a completed trial designed to assess the post-market risk margin of 1.4.

1.2 Statistical Issues and Post-Approval Implications

1.2.1 High Percentage of Treatment Discontinuations

At the time of the NDA resubmission, approximately 62% CONTRAVE subjects and 73% placebo subjects had discontinued treatment. These high percentages of treatment discontinuations call into question the ability to interpret the final results should the LIGHT trial continue to completion, given that the majority of events may be observed after the subjects have discontinued treatment; refer to Appendix 4 for more details.

1.2.2 Unblinding and LIGHT Trial Integrity

There are serious concerns raised in this review about the extent of unblinding and integrity of the ongoing LIGHT trial. As stated in the FDA Guidance for Clinical Trial Sponsors – *Establishment and Operation of Clinical Trial Data Monitoring Committees* (2006), "Knowledge of unblinded interim comparisons from a clinical trial is generally not necessary for those conducting or sponsoring the trial; further, such knowledge can bias the outcome of the

study by inappropriately influencing its continuing conduct or the plan of analyses." As such, unblinded interim results are expected to be shared only with the DMC and a select set of sponsor personnel under appropriate firewall. This group is also expected to be functionally independent of the operational aspects of the trial.

It is not clear if the necessary precautions were taken in the LIGHT trial, especially in light of the fact that over 100 individuals had unblinded knowledge of the interim findings, the majority of whom had "full access" to the interim data including board of directors and those with business interests in the trial outcome (refer to Section 3.1.2 and Appendix 5 for more details on the extent of unblinding). Individuals with full access are defined by the Applicant as "those that have access to unblinded, summarized, and individual subject study data."

This level of unblinding initiated by the Applicant raises serious concerns about the trial integrity. It is worth noting that there is no direct way to measure or assess the impacts of this level of unblinding on the conduct of the trial and its influence on the interpretability of data generated after the blind was broken. Due to the actions taken by the Applicant, it is questionable whether the integrity of the LIGHT trial is intact and consequently the reliability of the trial to rule out the 1.4 post-market risk margin.

As shown in Table 1, the findings from the planned pre-approval interim analysis of the LIGHT trial (b) (4) (b) (4) (b) (4) In the case of the LIGHT trial, the interim analysis is based on approximately 25% of the planned total number of MACE and the observed data may be a (b) (4) Thus, such an analysis is subject to the concerns raised above.

For serious risks, such as CV risk, that warrant assessment through a randomized clinical trial, it is imperative that the trial be conducted to the highest of scientific standards as there is typically only a single opportunity to reliably characterize the risk. In the original NDA submission, concerns were raised about the cardiovascular safety profile of CONTRAVE that warranted further investigation in a randomized cardiovascular outcome trial with the ultimate objective of ruling out a risk margin of 1.4. The LIGHT trial was adequately designed to achieve this objective. However, the Applicant has taken actions that have the potential to compromise the integrity of the LIGHT trial raising concerns about the ability to rely on data generated after the blind was broken to rule out the 1.4 risk margin. This in turn raises questions about the suitability of the LIGHT trial to achieve its ultimate objective in characterizing the CV risk of CONTRAVE. Due to these concerns one can postulate that the LIGHT trial is not being

conducted to the highest of scientific standards. As such, we recommend a new cardiovascular outcome trial that is held to the highest of scientific standards be initiated with the objective of ruling out a relative CV risk of 1.4.

2 INTRODUCTION

2.1 Overview and Regulatory Background

CONTRAVE is being developed by Orexigen Therapeutics Inc., the Applicant, as a fixed dose combination of naltrexone and bupropion, both of which are currently approved monotherapies in the United States. Naltrexone is approved for the treatment of opiate and alcohol dependence and bupropion is approved for the treatment of major depression and nicotine dependence. The New Drug Application (NDA 200063) for CONTRAVE was first submitted by the Applicant to the U.S. Food and Drug Administration (FDA) in March 2010 to be indicated for the treatment of obesity and weight management, including weight loss and maintenance of weight loss. The Applicant claims that the naltrexone/bupropion combination will interact in the central nervous system to decrease food intake and increase energy expenditure, resulting in weight loss. The proposed dose for CONTRAVE is two 8 mg naltrexone/90 mg bupropion tablets taken twice daily for a total dose of 32 mg/360 mg.

In January 2011, the Division of Metabolism and Endocrinology Products (DMEP) issued a Complete Response letter (CRL) citing concerns about the cardiovascular (CV) safety profile of CONTRAVE when used long-term in the intended population. These concerns were prompted by means for systolic and diastolic blood pressure and heart rate that were statistically significantly higher for CONTRAVE subjects compared to placebo subjects in the Phase 3 trials. Additionally, there were more adverse events related to hypertension in CONTRAVE subjects, especially those with type 2 diabetes. Therefore, the CRL stated that prior to approval of the NDA, the Applicant "must conduct a randomized, double-blind, placebo-controlled trial of sufficient size and duration to demonstrate that the risk of major adverse cardiovascular events in overweight and obese subjects treated with naltrexone/bupropion does not adversely affect the drug's benefit-risk profile"[8]. The remainder of this subsection summarizes the regulatory interaction with the Applicant pertaining to the development of the required CV outcomes trial (CVOT) and path to resubmission of the NDA for CONTRAVE.

An End of Review (EOR) meeting[9] was held with the Applicant, during which the discussion was focused on the Applicant's proposed design of the required CVOT. Note that a full protocol was not submitted for FDA review at the time of this meeting. During this meeting, FDA indicated that the appropriate primary endpoint for the trial is a composite endpoint, also known as MACE, comprising CV death, non-fatal myocardial infarction (MI), or non-fatal stroke. Additionally, the following points were noted:

[8] Refer to Complete Response letter dated January 31, 2011
[9] Refer to End of Review meeting minutes dated June 27, 2011

- FDA was concerned that the proposed annual background MACE rate of (b)(4)% would not include sufficient subjects at increased CV risk to adequately investigate the effect of CONTRAVE on CV risk. The recommendation was a background rate of between 1% and 1.5%.

- FDA stated that an interim analysis of at least 50% of MACE may be considered for approval provided the point estimate of the hazard ratio (HR) for MACE was close or below one and the upper bound of the confidence interval (CI) ruled out 10 excess events per 1000 person-years.

- FDA stated that the trial can be powered for the intent-to-treat (ITT) population, but that the results from the per-protocol (PP) population (also referred to as on-treatment results) need to be consistent and not divergent from the ITT results (also referred to as on-study results); see Section 3.3 for definition of on-study and on-treatment analyses. FDA also noted that powering the trial for only the ITT population may result in an underpowered trial for comparisons using the PP population.

Subsequent to EOR meeting, the Applicant submitted three Formal Dispute Resolution Requests (FDRRs) to the FDA all of which were denied. Among other issues, the Applicant queried whether the requirement for a pre-approval CVOT was consistent with recent FDA actions for similar products and noted that to meet the expectations discussed in the EOR meeting, the CVOT would need to enroll at least 100,000 subjects with trial duration between 3 to 4 years. The summaries of the FDRR decisions are provided below:

FDRR #1 to Office of Drug Evaluation 2 (ODE 2) dated June 8, 2011: This FDRR was denied by Director of the ODE 2, Dr. Curtis Rosebraugh in July 2011. In his denial letter, Dr. Rosebraugh noted that although weight reduction is undertaken for a variety of reasons, overweight and obesity have been widely associated with excess CV risk. Therefore, it stands contrary to the treatment of obesity to consider approval of a drug that may increase the risk of CV disease[10].

FDRR #2 dated July 14, 2011 to the Office of New Drugs (OND): This FDRR was denied by Director of OND, Dr. John Jenkins, in September 2011. In his denial letter, Dr. Jenkins outlined[11] broad design and analysis parameters for the CVOT that are provided below:

- The trial should enroll a population of overweight and obese subjects with an estimated background annual MACE rate of between 1.0% and 1.5%.

- The trial should follow subjects long enough to allow an assessment of long-term safety, that is, 2 to 3 years.

[10] Refer to FDRR Denial Letter by Dr. Curtis Rosebraugh, Director of ODE 2, dated July 7, 2011.
[11] Refer to FDRR Denial Letter by Dr. John Jenkins, Director of OND, dated September 15, 2011.

- The trial should be powered for the ITT analysis population; however, it does not require formal powering for the PP analysis population and the latter population will be assessed at the time of submittal.

- The pre-approval HR margin to be ruled out is 2.0 which would require at least 87 CV events. Post approval, the trial should rule out a HR of 1.4 requiring 371 CV events[12]. The HR margins are to be based on the upper bound of 95% (two-sided) confidence intervals (CIs).

FDRR #3 dated August 7, 2012 to the Center for Drug Evaluation and Research (CDER): This FDRR was denied by Dr. Douglas Throckmorton, Deputy Director for Regulatory Programs, CDER. In his denial letter[13], Dr. Throckmorton concurred with previous findings that there is sufficient uncertainty about the long-term CV safety of CONTRAVE that precluded approval.

In December 2011, prior to the FDRR #3, the Applicant submitted the original protocol for the CVOT with request for Special Protocol Assessment (SPA); SPA agreement dated February 3, 2012. The protocol, which followed the recommendations outlined in the FDRR #2 denial letter, was finalized on February 28, 2012 with subsequent amendments dated March 30, 2012 and January 28, 2013. The planned statistical analyses were included in the protocol submissions and reviewed by the FDA[14,15]; the Statistical Analysis Plan (SAP) was finalized on October 30, 2013. All CV events are to be prospectively adjudicated, in a blinded manner, by an independent Clinical Event Classification group (CEC), which is governed under a charter.

The pre-approval interim analysis of the required CVOT that is included in the resubmission for CONTRAVE, received December 11, 2013 (PDUFA Date: June 11, 2014), is the subject of this statistical safety review. The strategy for resubmission of the NDA was discussed with the Applicant at Type C meetings dated January 4, 2013; March 11, 2013 and October 24, 2013; and reflected in the respective meeting minutes[16,17,18]. As outlined in the denial letter to FDRR #2, the analysis of the interim data is conducted to rule out the pre-approval HR risk margin of 2.0. Note that the trial is ongoing at the time of this statistical review. According to the Applicant, the interim analysis and the data supporting resubmission was restricted to members of an unblinded team to ensure integrity of the ongoing CVOT. A discussion of the procedures implemented by the Applicant to protect the integrity of the trial data is provided in Section 3.1.

NOTE: The resubmission received December 11, 2013 included an interim analysis summary report, rather than the full interim clinical study report (CSR) which is usually required at the time of an NDA submission. Additionally, the submission did not include the final standardized datasets needed for this statistical review not included on this date. The FDA accepted an

[12] Calculations are based on 90% power using a log rank statistic assuming a HR = 1.0 and 1:1 randomization.
[13] Refer to FDRR Denial Letter by Dr. Douglas Throckmorton of OND, dated October 15, 2012.
[14] Refer to Statistical Review by Dr. Xiao Ding, DBVII, dated January 25, 2012.
[15] Refer to Statistical Review by Dr. Mat Soukup, DBVII, dated June 8, 2012.
[16] Refer to Type C meeting minutes dated January 31, 2013.
[17] Refer to Type C meeting minutes dated April 9, 2013.
[18] Refer to Type C meeting minutes dated November 12, 2013.

incomplete submission with the condition that the full interim analysis CSR and corresponding datasets must be submitted within 60 days[19] of the summary report per the recommendation of denial letter to FDRR #3. The Agency noted that failure of the Applicant to comply with this stipulation may result in Complete Response action or an extension of the review clock. The full study report and analysis datasets used in this review were submitted February 7, 2014; thereby meeting the submission timeline.

2.2 Data Sources

The NDA was resubmitted electronically and includes analysis datasets that are relevant for the interim analysis of the CVOT. All analysis datasets used for generating results in this statistical review were submitted by the Applicant in CDISC Analysis Dataset Model format. Detailed data definitions files for each of the analysis datasets are also included in the submission. Datasets and definition files can be found at the following location:

\\cdsesub1\evsprod\nda200063\0041\m5\datasets\nb-cvot\analysis\legacy\datasets

The following datasets were used in this statistical safety review:

- "adtte.xpt" which contains the time to CV event analysis variables
- "adsl.xpt" which contains the subject demographic and disposition variables
- "advs.xpt" which contains the vital signs variables
- "adae.xpt" which contains the adverse event variables

The quality and integrity of the data included in the submission will be discussed in Section 3.1.

3 STATISTICAL SAFETY EVALUATION

This is a statistical safety review that focuses on the CV safety assessment for CONTRAVE based on interim data from an ongoing CVOT, also known as the LIGHT trial. There are no efficacy evaluations performed in this review. Refer to clinical review by Dr. Eileen Craig for overall safety and efficacy evaluation of the CONTRAVE resubmission.

3.1 Data and Analysis Quality

3.1.1 Pre-Approval Findings: Quality of Interim Data and Interim Study Report

According to the study report, the CVOT is being conducted in accordance with ethical principles concerning medical research in humans that are consistent with Good Clinical Practice and applicable regulatory requirements. Additionally, the Applicant notes that quality control review of the interim CSR was performed to ensure consistency, clarity, and accuracy. There were no issues with the quality of the CSR identified during this statistical review.

[19] Refer to Type C meeting minutes dated January 31, 2013.

The data definition files provided sufficient details such that the primary endpoint analysis results could be replicated with ease from the submitted analysis datasets. There were no notable data quality or analysis issues discovered in this review that would impact the findings of the primary endpoint analyses for CV safety.

3.1.2 Post-Approval Implications: Unblinding and Trial Integrity

There is serious concern about the integrity of the ongoing LIGHT trial and consequently, the reliability of the final results of this trial to assess the post-market HR risk margin of 1.4. In the February 7, 2014 submission, the Applicant provided a Data Access Plan (DAP), which states that the unblinded team is limited to those individuals needed to facilitate or manage global regulatory submissions; refer to Table 2 for levels of data access of the unblinded team. Following request for information regarding the unblinded individuals, the Applicant provided[20] a list of individuals, excluding DMC members, who had knowledge of the interim results or access to unblinded interim data. There is serious concern about the extensive list of over 100 unblinded individuals, the majority of whom were granted full access to unblinded data of this ongoing trial (see Appendix 5 for the unblinded list submitted by the Applicant). According to the Applicant, individuals with full access are defined as "those that have access to unblinded, summarized, and individual subject study data". It is particularly concerning that members of Orexigen's Board of Directors who have financial interest in the outcome of the trial were also provided full access to unblinded data. Furthermore, there is concern that the unblinded list provided by the Applicant does not include the names of potential (b) (4) who have obtained knowledge of the unblinded interim results. According the Applicant, the names of these potential (b) (4) were not provided because "we [Orexigen] have entered into confidentiality agreements with each of the potential (b) (4) .

As stated in the FDA Guidance for Clinical Trial Sponsors – *Establishment and Operation of Clinical Trial Data Monitoring Committees* (2006), "Knowledge of unblinded interim comparisons from a clinical trial is generally not necessary for those conducting or sponsoring the trial; further, such knowledge can bias the outcome of the study by inappropriately influencing its continuing conduct or the plan of analyses." As such, unblinded interim results are expected to be shared only with the DMC and a select set of sponsor personnel under appropriate firewall. This group is also expected to be functionally independent of the operational aspects of the trial. It is not clear if the necessary precautions were taken in the ongoing LIGHT trial, especially in light of the fact that over 100 individuals had unblinded knowledge of the interim findings or access to unblinded data. This extent of unblinding initiated by the Applicant raises serious concerns about the ongoing trial integrity. It is worth noting that there is no direct way to measure or assess the impacts of this level of unblinding on the conduct of the trial and its influence on the interpretability of data generated after the blind was broken. Due to the actions taken by the Applicant, it is questionable whether the integrity of the LIGHT

[20] Refer to Applicant's response to information request dated April 16, 2014.

trial is intact and consequently the reliability of the trial to rule out the 1.4 post-market risk margin.

Table 2 Levels of Access to Unblinded Data for CVOT

Access Level	Description
Full Access	Individuals and entities with full access to unblinded data from the 87-event interim analysis are defined as those that have access to unblinded, summarized and individual subject study data. This will generally be in the form of tables, figures, and listings containing unblinded study data (e.g., interim analysis DMC Closed Report).
Knowledge of Threshold	Individuals and entities informed at this level will be advised prior to the press release whether the DMC, based on careful review of available evidence and in the context of pre-specified monitoring boundaries, has recommended the 87-event interim data should be released to the core group of individuals essential to the facilitation of regulatory submissions. Individuals with knowledge of this information prior to the press release are considered to be in possession of confidential information up to the point of the press release, which is planned to contain the same level of information.

Source: Extracted from Data Access Plan Table 1 (page 5)

3.2 Study Design and Endpoints

3.2.1 Study Design

The CVOT is an ongoing event-driven, phase 3b, multicenter, randomized, double-blind, placebo-controlled trial to assess the occurrence of major CV adverse events in overweight and obese subjects with $27kg/m^2 \leq BMI \leq 50kg/m^2$ at increased risk of adverse CV outcomes receiving CONTRAVE relative to placebo. According to the protocol, subjects at increased risk of adverse CV outcomes are those with at least one of the following 2 conditions:

1. CV disease (confirmed or at high likelihood of CV disease) with at least one of the following:

    - History of documented MI at least 3 months prior to screening
    - History of coronary revascularization
    - History of carotid or peripheral revascularization
    - Angina with ischemic changes or positive cardiac imaging study
    - Ankle brachial index<0.9 within prior 2 years
    - At least 50% stenosis of a coronary, carotid, or lower extremity artery within prior 2 years

2. Type 2 diabetes mellitus (T2DM) with at least 2 of the following:

    - Concurrent hypertension

- Dyslipidemia currently treated with FDA-approved pharmacotherapy or documented high LDL cholesterol (>100 mg/dL) within 12 months
- Documented low HDL cholesterol(<50 mg/dL in women or <40 mg/dL in men) within 12 months
- Current tobacco smoker

To be eligible for trial participation, men must be at least 50 years of age and women at least 45 years of age. Additional inclusion/exclusion criteria are provided in Section 4 of the protocol.

The trial is being conducted at 264 sites in the United States. Approximately 10500 subjects were to be enrolled in the trial, see Section 3.3.2 for details of the Applicant's sample size estimation. The trial has 3 periods which are conducted in the following sequence: a screening period of up to 2 weeks to verify eligibility, a 2-week double-blind lead-in period, and a double-blind treatment period of approximately 208 weeks. During the lead-in period subjects are randomized in a 1:1 ratio via a centralized Interactive Voice or Web Response System (IVRS/IWRS) to receive treatment according to one of two treatment sequences: 1 week of CONTRAVE with dosage 8 mg naltrexone sustained release (SR)/90 mg bupropion SR tablet taken once daily followed by one week of once daily matching placebo, or vice versa. During the treatment period eligible subjects who complete the lead-in period and do not have a suspected CV event, are randomized in a 1:1 ratio via IVRS/IWRS to either CONTRAVE or matching placebo. No stratification factors are used in the randomization process. A summary of the design and visit schedule for the trial is shown in Figure 1. During the first 4 weeks of the treatment period, CONTRAVE is titrated to a maximum dosage of 32 naltrexone SR/360 mg bupropion SR tablet, see Figure 2.

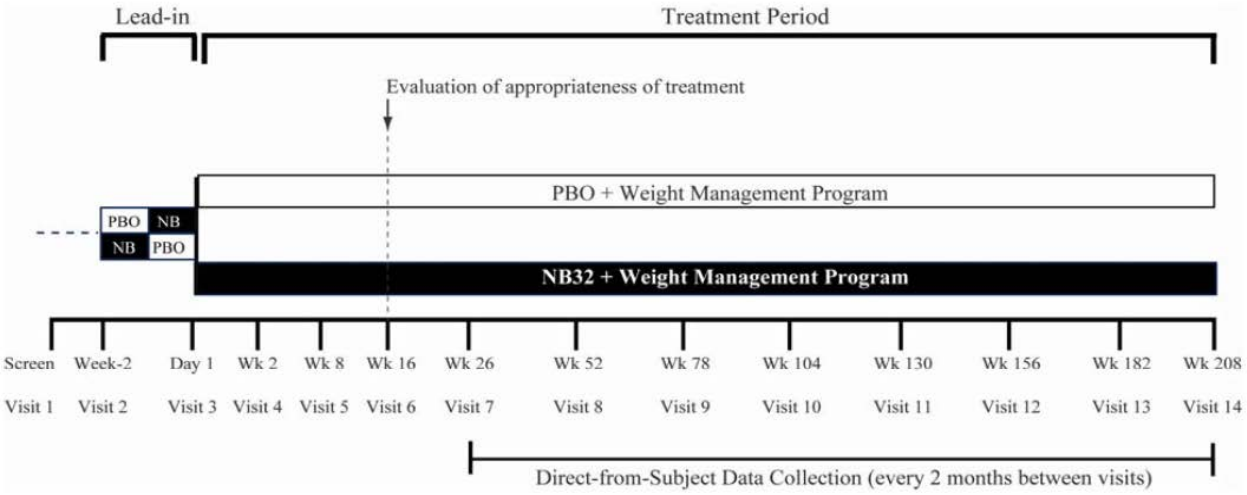
Baseline measurements including body weight and vital signs, queries for serious adverse events, and registration in the weight management program, are conducted at Visit 3 (Day 1 of follow-up) prior to subject randomization. All randomized subjects continue routine medical care, including management of diabetes and other comorbid conditions from their usual health care provider, which may or may not be the trial Investigator. Subjects return for study site visits according to the schedule shown in Figure 1. At Week 16 (Visit 6), there is a planned evaluation of weight loss and blood pressure relative to baseline measurements. Subjects are discontinued from trial medication at this visit if:

- They have not lost at least 2% of their body weight or
- They are experiencing consecutive, sustained increases in systolic or diastolic blood pressure of ≥10 mm Hg.

According to the study report, these Week 16 evaluation thresholds are based on retrospective analyses of data from phase 3 trials to assess the impact of various degrees of weight loss and blood pressure increases in determining which subjects are most likely to respond favorably to continued treatment. Subjects continue to participate in the weight management program while enrolled in the trial, regardless of whether they continue taking trial medication. Starting after Week 26 (Visit 7) and continuing every other month between site visits, all subjects are required

to access an internet- or telephone-based assessment tool to record hospitalizations and medication compliance.

Figure 1 CVOT Design and Visit Schedule



Source: Extracted from CVOT Protocol Figure 1 (page 18)

Figure 2 CVOT Dosing Schedule

Dose Schedule	Lead-in Period		Treatment Period			
	Week -2	Week -1	Week 1 (Days 1-7)	Week 2 (Days 8-14)	Week 3 (Days 15-21)	Week 4 to end of study
*Total Daily Dose**	*8/90 NB*	*8/90 NB*	*8/90 NB*	*16/180 NB*	*24/270 NB*	*32/360 NB*
Morning	1 tab NB or PBO	1 tab PBO or NB	1 tab NB or PBO	1 tab NB or PBO	2 tabs NB or PBO	2 tabs NB or PBO
Evening	--	--	--	1 tab NB or PBO	1 tab NB or PBO	2 tabs NB or PBO

*Doses shown are of naltrexone SR/bupropion SR (NB); tab=tablet; PBO=placebo.

Source: Extracted from CVOT Protocol Table 1 (page 27)

At any time during the trial subjects may choose to discontinue trial medication prematurely or may be asked to discontinue medication by the Investigator because of noncompliance with medication or procedures, or adverse event. All subjects who prematurely discontinue trial medication are to complete end of treatment visit procedures and continue to participate in the trial for collection of safety data through trial completion. Subjects who complete the end of treatment visit procedures may not reinitiate taking trial medication at any time during the trial. Subjects who discontinue from study procedures, that is, weight management program, study site visits, and internet- or telephone-based contacts, will be contacted every 6 months or a minimum of one contact prior to trial completion to collect information on potential primary CV events, provided the subject has not revoked consent for all further follow-up. Refer to Section 3.4 for discussion of subject disposition.

There are three committees responsible for safety oversight of the trial: Executive Steering Committee (ESC), Clinical Endpoint Committee (CEC) and Data Monitoring Committee (DMC). The ESC, composed of experts in CVOTs, cardiology, obesity and endocrinology, provides strategic and scientific advice to ensure the most appropriate conduct and execution of the trial. The CEC, composed of an independent panel of cardiology and neurology experts, adjudicates the CV events in a consistent manner; see Section 3.2.2 for details of the adjudication process. According to the study report, the ESC and CEC remain blinded to randomized treatment assignment and have no knowledge of unblinded aggregate or subject level data at any time during the trial.

The DMC is responsible for monitoring the subject safety during the trial and providing recommendations about stopping or continuing the trial. The DMC is composed of individuals external to the trial organizers, Applicant, and Investigators, and operates under a written charter (dated August 7, 2013) that is included in the resubmission. The DMC holds regular meetings using Open and Closed session formats to preserve confidentiality of the ongoing CVOT while at the same time providing opportunities for interaction between the DMC and Applicant. The reports for these sessions are prepared by (b) (4) which is an unblinded independent statistical center. The Open session reports include data on recruitment, baseline characteristics, eligibility violations, completeness of follow-up, and compliance for all subjects randomized into the double-blind treatment period, without specification of treatment group. The Closed session reports include analyses of primary, secondary, and additional endpoints, subgroups analyses, as well as the summaries provided in the Open session reports. All analyses included in the Closed session report are presented by coded treatment group for all subjects randomized into the double-blind treatment period; the DMC members are provided the treatment codes.

Reviewer's Comment: The DMC reviewed the results of the interim analysis that is the subject of this statistical review and informed the Applicant that the pre-approval[21] threshold of ruling out a HR of 2.0 had been met. Therefore, as agreed upon[22] with the FDA, the DMC Open and Closed reports served as the basis for the resubmission of the NDA received December 11, 2013.

3.2.2 Definition of Endpoints and Adjudication Process

The primary endpoint of the trial also referred to as **MACE** throughout this review, is a composite comprising CV death (including fatal MI and stroke), nonfatal MI, or nonfatal stroke. The secondary endpoint also referred to as **MACE+** throughout this review, is a composite of CV death, nonfatal MI, nonfatal stroke, or nonfatal unstable angina requiring hospitalization. All-cause mortality is also assessed in this review.

All suspected CV events are adjudicated by an independent CEC from the Cleveland Clinic Center for Clinical Research. The CEC members include 3 cardiologists and 3 neurologists and are governed under a charter (finalized February 24, 2012). The adjudication is being

[21] Refer to FDRR Denial Letter by Dr. John Jenkins, Director of OND dated September 15, 2011
[22] Refer to Type C Meeting minutes dated January 31, 2013

prospectively conducted in a blinded consistent manner using standardized definitions[23]. Suspected events are identified for adjudication in the following ways:

- Events identified by the Investigator during regular assessments of SAEs.
- Events discovered by the site monitor at monitoring visits that had not been previously reported.
- Events identified during periodic database queries for preferred terms that are triggers for adjudication (e.g. the preferred term of apraxia triggers a stroke review by the CEC).
- Events identified during review of source documents by the CEC.

Once a site has knowledge of a suspected CV event, the designated site personnel enters the appropriate subject information needed for adjudication into the electronic data capture system. The minimum information required for adjudication is a subject package containing supporting documentation (e.g. summary of hospitalization or death certificate) and either a narrative completed by the Investigator or complete source documents.

The adjudication of each suspected CV event is performed in two phases. The initial phase entails independent review of subject package by two adjudicators. Each adjudicated event is classified as a confirmed event (meeting the event definition with all necessary documentation), a non-event (does not meet the event definition and likely represents an alternative or non-event diagnosis), or lack of documentation for confirmation of an event. If the independent adjudicators agree at this phase, the adjudication is considered final. If there is disagreement during the initial review phase, the package is forwarded to the second phase, where it is reviewed by the CEC Director, senior cardiologist for CV events, or senior neurologist for neurology events. The second phase of the adjudication is considered final, and supersedes any prior decision.

3.3 Statistical Methodologies

This section describes the primary analysis population, statistical hypotheses, planned analyses and other statistical components outlined in the SAP, dated October 30, 2013, that are relevant for this statistical review. The Applicant notes that all analyses included in the SAP are considered *a priori* as they were specified before unblinding of the data and cut-off date for the interim analysis performed on November 6, 2013. All other analyses that are included in the study report that are not described in this SAP are considered post hoc. This section also describes additional analyses conducted by the statistical reviewer.

The primary statistic discussed in this review is the hazard ratio (HR), CONTRAVE relative to placebo, for the endpoints defined in Section 3.2.2. A HR of one is indicative of equivalent rates

[23] Refer to *Standardized Definitions for Endpoint Events in Cardiovascular Trials* dated November 9, 2012

between the two treatment groups, a HR greater than one is indicative of higher rate in the CONTRAVE compared to placebo, and vice versa for HR less than one.

3.3.1 Analysis Population

The primary analysis population, also referred to as intent-to-treat (ITT) population, includes all subjects who undergo randomization into the double-blind treatment period and are dispensed trial medication, where dispensed is defined as having at least one medical ID number on the medication dispensation and administration case report form. This population is used for all analyses presented in this review.

Reviewer's Comment: The Applicant defines a per protocol (PP) population for sensitivity analyses, such as the on-treatment analysis, that is composed of ITT subjects who take at least one dose of trial medication during the treatment period in accordance with the trial protocol. This PP population is a subset of the ITT population as it contains only those subjects who are treated in accordance with the trial protocol; therefore, analyses based on the PP population are excluded from this review.

3.3.2 Statistical Hypotheses and Sample Size Estimation

The following are the pre-specified statistical hypotheses to be tested by trial completion with respect to the primary endpoint:

H01: HR for CONTRAVE relative to placebo ≥2.0 (required pre-approval after 87 MACE)
H02: HR for CONTRAVE relative to placebo ≥1.4 (required post-approval after 371 MACE)
(b) (4)

Note that H01 and H02 were stipulated in the denial letter[24] to FDRR#2; however, this statistical review is based only on hypothesis H01, which assesses the pre-approval threshold. Testing of H01 uses its own significance level of alpha=0.025 (one-sided); therefore, assessment of ruling out the 2.0 risk margin is based on the two-sided 95% confidence interval.

The Applicant's power and sample size calculations were based on the post-approval risk margin of 1.4. The following assumptions were used to determine the number of MACE required:

- True HR of 1.0
- One-sided alpha of 0.025
- Annual background MACE rate of 1.5%
- 2 interim analyses (at 50% and 75% MACE), assuming O'Brien Fleming spending function to control alpha level

Under these assumptions, a total of 378 MACE are needed to achieve 90% power to rule out HR of 1.4 post approval. The Applicant further estimated that to obtain this many events, the trial

[24] Refer to FDRR Denial Letter by Dr. John Jenkins, Director of OND dated September 15, 2011

would need approximately 4450 subjects per treatment arm with the additional assumptions of 1:1 randomization, 0.5 year recruitment, maximum subject follow-up of 4 years, and annual lost to follow-up rate of 0.012.

Reviewer's Comments:

1. ***The Applicant outlined a testing strategy for maintaining the Type I error rate at 0.025 (one-sided) assuming two planned interim analyses for testing H02 (b) (4) (b) (4). There were no plans for testing H02 (b) (4) at the time that H01 is tested. Therefore, no conclusions about the 1.4 risk margin can be made from this interim data.***

2. ***Note that the Applicant's estimated total number of events of 378 to rule out 1.4 takes into account plans to conduct two interim analyses. This assumption was not made in the denial letter to FDRR#2, which stipulated that 371 events are needed.***

3.3.3 Applicant's Planned Statistical Analyses

The agreed upon primary statistical analysis is time from randomization to first occurrence of positively adjudicated MACE. All MACE events, irrespective of subject's treatment status, that occurred after randomization and prior to the interim analysis data cut-off date of November 6, 2013 are included in this analysis. This is referred to as the on-study analysis in this review. For each subject with no event, the time at risk is the duration of time on study after randomization. For each subject with an event, the time at risk is the time from randomization to the first occurrence of an outcome comprising MACE. The HR for MACE, and corresponding 95% CI, are estimated using a Cox proportional hazards (PH) model with treatment as the only factor.

Similar Cox PH models are used to estimate HRs and 95% CIs for the secondary endpoint (MACE+), individual MACE component, and all-cause mortality. Additionally, Cox PH models are used for subgroup analyses for age, sex, race, and CV risk factors.

Separate incidence plots of MACE and MACE+ over time, by treatment group, are plotted using Kaplan-Meier plots.

Reviewer's Comment: There are no changes to the pre-specified primary and secondary analyses that were identified in the report.

3.3.4 Reviewer's Post Hoc Analyses

Sensitivity analysis, also referred to as on-treatment analysis, of MACE are performed utilizing all subjects in the ITT population and including only those events occurring while the subjects are still on treatment or within 30 days of last treatment dose. A similar Cox PH model as used for the primary analysis is used to estimate the HR and 95% CI for the on-treatment analysis.

Sensitivity analyses are also conducted to evaluate the possibility of a biased result due to informative censoring in the on-study and on-treatment analyses of MACE. In general, informative censoring arises when censoring of subjects is related to the chance of the event occurring. In this trial, informative censoring could occur if censored subjects who have SAEs that could have later developed into MACE events were discontinued from follow-up shortly after the SAE occurred. Two informative censoring exploratory analyses are conducted:

- On-study informative censoring: This analysis includes subjects who were originally censored due to lost to follow up for MACE. Subjects with at least one cardiac SAE or SAE with preferred term including "stroke" reported in the adverse event dataset, which occurred (started or ended) within 30 days of the last contact date are included as MACE in the analysis.

- On-treatment informative censoring: This analysis includes subjects originally censored in the on-treatment analysis due to treatment discontinuation. Subjects with at least one cardiac SAE or SAE with preferred term including "stroke" reported in the adverse event dataset, which occurred (started or ended) within 30 days of the last treatment date are included as MACE in this analysis.

In these sensitivity analyses, for any subject with a missing adverse event end date, the assumption is made that the event occurred within the window specified above, regardless if the start date of the AE was more than 30 days before the end of follow-up window. For all subjects with events meeting the criteria defined above, the time at risk for MACE is duration from randomization date to AE onset date.

Reviewer's Comments: It is important to note that these sensitivity analyses are not based on type of SAE or potential for MACE development from a clinical perspective, but rather timing of SAE occurrence. Additionally, these analyses include events that might have been adjudicated as MACE+.

3.4 Subject Disposition, Demographic and Baseline Characteristics

There were 13192 subjects screened for trial participation; of which 10514 were randomized for the lead-in period. There were 1500 subjects who did not complete the lead-in period, primarily due to adverse events or not meeting inclusion/exclusion criteria. Among the 9014 subjects who completed the lead-in period, 104 subjects were not randomized into the double-blind treatment period. Therefore, 8910 subjects were randomized (4456 to CONTRAVE and 4454 placebo), of which 5 subjects were not dispensed study medication. As a result, the ITT population used for all analyses in this review comprised 8905 subjects: 4455 randomized to CONTRAVE and 4450 randomized to placebo. Note that at the time of this review, enrollment into the trial was complete; therefore, the distributions of baseline characteristics that follow should be the same at the end of the trial. However, changes are expected to the disposition of subjects at the end of the trial as subjects are still being followed.

As of the interim analysis cut-off date, there are 28.6% CONTRAVE subjects and 31.5% placebo subjects who have discontinued the trial; see Table 3 for summary of reasons for trial discontinuation. In accordance with the protocol, some subjects who discontinued the trial are continued to be followed (e.g. through contact with subject's primary physician) for MACE events. The percentages of subjects who were completely loss to follow-up for MACE events were similar between the two treatment groups (4.9% for CONTRAVE and 4.7% for placebo). Therefore, approximately 95% of the ITT population was retained for the MACE analyses.

***Reviewer's Comment**: The annual rate of discontinuation assumed at the design stage of the trial was 1.2%. As of the interim analysis, which is approximately 1.5 years into the trial, the lost to follow-up rate was approximately 5%. Should the trial continue, this higher lost to follow-up rate could potentially impact the ability of the trial to attain the expected number of events needed by the estimated maximum duration of 4 years. The DMC recommended[25] that the "study teams should continue to be pro-active in identifying and implementing creative approaches for reducing the fraction of subjects who are missing MACE follow-up." The Applicant has collected vital signs data from public records for approximately 80% of those subjects who were lost to follow-up for MACE. It is unclear if public records can be used for uncovering MACE outcome data.*

Table 3 Distribution of Subject Trial Status by Treatment Group

Subject Status	CONTRAVE (N=4455) n (%)	Placebo (N=4450) n (%)
Continuing on trial	**3178 (71.4)**	**3047 (68.5)**
On treatment	1709 (38.4)	1201 (27.0)
Discontinued treatment	1469 (33.0)	1846 (41.5)
Discontinued trial	**1277 (28.6)**	**1403 (31.5)**
Discontinued trial procedures[1]	1006 (22.6)	1124 (25.3)
Discontinued MACE follow-up[2]	52 (1.2)	70 (1.6)
Discontinued any follow-up contact[3]	102 (2.3)	90 (2.0)
Lost to follow-up	117 (2.6)	119 (2.7)

Shaded region indicates subjects who are completely loss to follow-up for MACE
[1]Subjects discontinued procedures, e.g. site visits and weight management program, but are contacted periodically (every 6 months or at a minimum a single contact prior to trial completion) for potential MACE events.
[2] Subjects are unwilling to be contacted on periodic basis for MACE follow-up, but the subject's primary physician or designated contact are contacted periodically for potential MACE.
[3]Subjects revoke consent to contact their primary physician or designated contact for MACE follow-up.
Source: Created by the statistical reviewer using dataset "adsl.xpt"

There are 2732/4455 (61.6%) CONTRAVE subjects and 3248/4450 (73%) placebo subjects who discontinued treatment; see Table 4 for summary of reasons for treatment discontinuation. Note that this table includes subjects who discontinued treatment but remained in the trial. The most common reason for treatment discontinuation in CONTRAVE subjects was due to adverse events (26.7% CONTRAVE, 7.4% placebo). According to the study report, the AEs in the

[25] Refer to DMC meeting minutes for meeting held November 23, 2013

CONTRAVE arm are primarily tolerability related, e.g. nausea. The most common reason for treatment discontinuation in placebo subjects is not meeting week 16 criteria (14.2% CONTRAVE, 40.7% placebo); primarily due to not achieving at least 2% weight loss as stipulated in the protocol.

Table 4 Summary of Reasons for Treatment Discontinuation

Reason for Treatment Discontinuation	CONTRAVE (N=4455) n (%)	Placebo (N=4450) n (%)
Adverse event	1188 (26.7)	330 (7.4)
Subject decision	613 (13.8)	778 (17.5)
Protocol deviation	14 (0.3)	27 (0.6)
Lost to follow-up	158 (3.6)	162 (3.6)
Sponsor decision	14 (0.3)	25 (0.6)
Not meeting week 16 criteria*:	633 (14.2)	1809 (40.7)
Weight only	471 (10.6)	1559 (35.0)
BP only	59 (1.3)	17 (0.4)
Weight and BP	96 (2.2)	213 (4.8)
Other**	7 (0.2)	20 (0.4)
Other	126 (2.8)	117 (2.6)
Total	2746 (61.6)	3248 (73.0)

BP=blood pressure
*Two consecutive, sustained (at week 8 and week 16 or week 2 and week 16 week 8 is missing) increases in systolic or diastolic blood pressure of ≥10 mm Hg or weight loss not less than 2% of baseline.
**27 subjects (7 CONTRAVE subjects, 20 placebo subjects) recorded as 'Other' due to incomplete data, inconsistent data, or incorrect eCRF entry.

Source: Created by the statistical reviewer using dataset "adsl.xpt" and "advs.xpt"

The distributions of follow-up for MACE is similar between the CONTRAVE and placebo subjects: 56.2 weeks (range: 0.42 – 72.4) for CONTRAVE and 56.2 weeks (range: 0.14 – 73.1) for placebo. The extent of exposure to treatment is longer in the CONTRAVE subjects (mean: 30.5 weeks, range: 0.14 – 72.4) compared to placebo subjects (mean: 26.8 weeks, range: 0.14 – 72.4).

The observed medication retention rates, or rates for subjects who have not discontinued treatment, are 38.4% in CONTRAVE subjects and 27% in placebo subjects. These rates are significantly lower than what was predicted by the Applicant for the 1.5 year time point, that is, the approximate time point of the interim analysis, and more consistent with the predicted rates at Year 3; see Table 5.

Table 5 Applicant's Projection of Cumulative Medication Retention Rates

Time Point	Placebo	NB32
Week 16**	51	56
Year 1	39	48
Year 2	32	40
Year 3	28	36
Year 4	27	33

NB32=CONTRAVE
Source: Extracted from the briefing package for Type A meeting held November 16, 2011

***Reviewer's Comment*: Given the high treatment discontinuation rates, it is expected that few events will be observed while subjects are still on treatment, which may result in wide confidence intervals for the on-treatment analysis; refer to Appendix 4 of this review for more discussion of this issue.**

Table 6 shows similar demographic characteristics for the treatment groups. The majority of subjects are female (55%) or White (84%). The average age of subjects is 61 years and average BMI is 37 kg/m^2. Recall that this trial was conducted in sites in the US only; therefore, treatment distributions for geographic region are not applicable in this review.

Table 7 shows similar distributions for baseline CV risk factors between the treatment groups. The majority of subjects had T2DM (85%) and a minority of subjects had a history of CV disease (32%). Very few subjects had a history of tobacco use (approximately 9%) and most subjects had normal renal function.

Table 6 Distribution of Demographic Characteristics by Treatment (ITT Population)

Demographic Characteristic	CONTRAVE N=4455 n (%)	Placebo N=4450 n (%)
Sex		
Male	2018 (45.3)	2031 (45.6)
Female	2437 (54.7)	2419 (54.4)
Age Category, n (%)		
<65	2973 (66.7)	3053 (68.6)
≥65	1482 (33.3)	1397 (31.4)
Age, in years		
Mean (SD)	61.1 (7.3)	60.9 (7.4)
Range	45 – 86	45 – 85
Race		
White	3738 (83.9)	3698 (83.1)
Non-white*	716 (16.1)	750 (16.9)
BMI Category, n (%)		
<35	1691 (38.0)	1719 (38.6)
35 – 40	1477 (33.2)	1383 (31.1)
≥40	1285 (48.8)	1348 (30.3)
BMI, in kg/m^2		
Mean (SD)	37.2 (5.3)	37.4 (5.4)
Range	27.0 – 50.4	26.6 – 50.8

SD=standard deviation
*Non-white includes American Indian/Alaska native, Asian, Black/African American, Hawaiian/Pacific Islander or other.

Source: Created by the statistical reviewer using dataset "adsl.xpt"

Table 7 Distribution of Baseline CV Risk Factors by Treatment (ITT Population)

Risk Factor	CONTRAVE N=4455 n (%)	Placebo N=4450) n (%)
History of CVD*	1414 (31.7)	1447 (32.5)
History of hypertension	4160 (93.4)	4114 (92.5)
History of dyslipidemia	4102 (92.1)	4070 (91.5)
History of tobacco use	405 (9.1)	414 (9.3)
History of T2DM	3783 (84.9)	3803 (85.5)
Duration of Diabetes**		
< 6 years	1494 (33.5)	1561 (35.1)
≥ 6 years	2205 (49.5)	2166 (48.7)
Renal Function, eGFR		
<90	1220 (27.4)	1174 (26.4)
≥90	3234 (72.6)	3275 (73.6)

CVD=cardiovascular disease, T2DM=type 2 diabetes mellitus

*History of any of the following: myocardial infarction, coronary artery bypass graft surgery, carotid endaterectomy, stent replacement, PTCA, lacer atherectomy, peripheral revascularization, lower leg atheroscelerotic atherectomy, repair of abdominal aotor aneurysm, femoral bypass, popliteal bypass, coronary revascularization, carotid revascularization, angina, ankle brachial index<0.9, ≥50% coronary stenosis in prior 2 years, ≥50% carotid stenosis in prior 2 years, ≥50% lower extremity artery in prior 2 years
** Duration of diabetes summarized only for subjects reported to have history of diabetes at baseline. There were 160 subjects (84 CONTRAVE, 76 placebo) with history of diabetes at baseline that were missing diabetes onset date and so duration could not be determined.

Source: Created by the statistical reviewer using dataset "adsl.xpt"

3.5 Results of Statistical Analyses

3.5.1 Results of MACE Analyses

Results of Primary (On-study) MACE Analysis

As of the pre-approval interim analysis, the incidence of MACE is (b) (4) (b) (4) The estimated HR is (b) (4) No major concerns about violation of the proportional hazards assumption were seen, see Appendix 1. As shown in Figure 3, subjects in the CONTRAVE group are (b) (4) Note that because the CVOT is ongoing, the numbers of subjects at risk shown in this plot are indicative of the subjects' follow-up time observed at the time of the interim analysis rather than trial discontinuations.

5 Page(s) has been Withheld in Full as b4 (CCI/TS) immediately following this page

5 SUMMARY AND CONCLUSIONS

5.1 Collective Evidence and Statistical Issues

The ongoing CV outcomes trial (CVOT), also known as the LIGHT trial, is being conducted in accordance with a protocol that was reviewed and agreed upon under SPA[26] with FDA. The objective of the trial is to demonstrate that CONTRAVE, a proposed product for weight management, is not associated with an increased risk of major adverse CV events in comparison to placebo when used in high risk subjects. The agreed upon primary analysis population comprises all randomized subjects who are dispensed study medication. The primary safety endpoint is MACE, a composite endpoint comprising cardiovascular death, non-fatal myocardial infarction or non-fatal stoke. A key secondary endpoint is MACE+, a composite endpoint comprising cardiovascular death, non-fatal myocardial infarction, non-fatal stoke, or hospitalization for unstable angina. All CV events included in analyses are prospectively adjudicated, in a blinded manner, by an independent Clinical Event Classification group, which is governed under a charter using standardized definitions. This statistical review is based on a pre-approval[27] interim analysis of this ongoing CVOT to show that a HR (CONTRAVE to placebo) risk margin of 2.0 can be ruled out; post-approval the trial is planned to rule out a HR risk margin of 1.4. The pre-specified primary statistical analysis uses a Cox proportional hazards model and included all events, regardless of subject's treatment status, that occurred after randomization through interim analysis data cut-off date. On-treatment analysis of MACE is also conducted which includes only those events that occur while subjects are still being treatment and within 30 days after treatment discontinuation. There are no adjustments for multiple testing of the interim data.

5.1.1 Collective Evidence from Pre-Approval Interim Analysis

As of the pre-approval interim analysis (at approximately 1.5 years of trial follow-up), the incidence of positively adjudicated MACE is (b) (4) (b) (4) The estimated hazard ratio for MACE is (b) (4) (b) (4) The incidence of MACE while subjects are on-treatment or within 30 days of treatment discontinuation is (b) (4) and the upper bound of the 95% CI is less than 2.0. Table 11 summarizes the analysis results for MACE, MACE+, and other endpoints assessed in this review. Note that the analyses of individual MACE components are meant for descriptive purposes only: (b) (4)

[26] Refer to Special Protocol Agreement letter dated February 3, 2012.
[27] Refer to FDR Denial Letter by Dr. John Jenkins, Director of OND dated September 15, 2011.

Table 11 Summary of Interim Analysis Results

Endpoint	CONTRAVE N=4455 n (%)	Placebo N=4450 n (%)	HR (95% CI)
MACE			
On-study[1]		(b) (4)	
On-treatment[2]			
MACE+			
MACE Components			
CV Death			
Nonfatal Stroke			
Nonfatal MI			
All-Cause Mortality			

n=number of subjects with event, N=number of subjects randomized, CI=confidence interval, HR=hazard ratio
[1] Primary on-study analysis includes all MACE that occurred after randomization up to interim analysis cut-off date, regardless of treatment exposure
[2] Sensitivity on-treatment analysis includes all MACE that occurred after randomization and within 30 days of treatment discontinuation

Source: Created by the statistical reviewer using dataset "adtte.xpt"

5.1.2 Statistical Issues and Post-Approval Implications

An important issue is the interpretability of the final results of the LIGHT trial should it continue to completion. The Applicant presents conditional power calculations that suggest, given the interim findings, the trial will have (b) (4) power to rule out the post-approval risk margin of 1.4. Although no issues are noted with these calculations, the concern is the impact of the high treatment discontinuation rates. There are 62% CONTRAVE subjects and 73% placebo subjects who have discontinued treatment as of the pre-approval interim analysis, which is approximately 1.5 years into the trial with planned maximum duration of 4 years. These high percentages of treatment discontinuations call into question the ability to interpret the results at end of the trial, given that the majority of events may be observed after the subjects have discontinued treatment; refer to Appendix 4.

More importantly, given the extent of unblinding in this ongoing trial, there is serious concern that the integrity of this trial has been compromised. As described in Section 3.1 and shown in Appendix 5, there are over 100 individuals with unblinded knowledge of the interim findings, the majority of whom had "full access" to the interim data, including board of directors and those with business interests in the trial outcome. Individuals with full access are defined by the

Applicant as "those that have access to unblinded, summarized, and individual subject study data." *Clinical Trial Data Monitoring Committees* (2006), "Knowledge of unblinded interim comparisons from a clinical trial is generally not necessary for those conducting or sponsoring the trial; further, such knowledge can bias the outcome of the study by inappropriately influencing its continuing conduct or the plan of analyses." As such, unblinded interim results are expected to be shared only with the DMC and a select set of sponsor personnel under appropriate firewall. This group is also expected to be functionally independent of the operational aspects of the trial. It is not clear if the necessary precautions were taken in the ongoing LIGHT trial. This level of unblinding initiated by the Applicant raises serious concerns about the trial integrity. It is worth noting that there is no direct way to measure or assess the impacts of this level of unblinding on the conduct of the trial and its influence on the interpretability of data generated after the blind was broken. Due to the actions taken on part of the Applicant, it is questionable whether the integrity of the LIGHT trial is intact and consequently the reliability of the trial to rule out the 1.4 post-market risk margin.

The findings from the pre-approval interim analysis of the LIGHT trial (b) (4) (b) (4) In the case of the LIGHT trial, the pre-approval interim analysis is based on approximately 25% of the planned total number of MACE and the observed data may be (b) (4) Thus, such an analysis is subject to the concerns raised above.

5.2 Conclusions and Recommendations

This is a statistical safety review of a pre-approval interim analysis, submitted by Orexigen Therapeutics Inc., the Applicant for this NDA, to assess CV safety of CONTRAVE compared to placebo in high risk subjects. The estimated HR for MACE (the primary endpoint) is (b) (4) The upper bound of this confidence interval is less than 2.0 and therefore ruled out the pre-approval risk margin of 2.0 set forth by the FDA.

While the conclusion is that pre-approval threshold has been met, the concern is that the extent of unblinding limits the ability of LIGHT trial data accrued after the blind was broken to assess the post-market risk margin. For serious risks, such as CV risk, that warrant assessment through a randomized clinical trial, it is imperative that the trial be conducted to the highest of scientific standards as there is typically only a single opportunity to reliably characterize the risk. In the

original NDA submission, concerns[28] were raised about the cardiovascular safety profile of CONTRAVE that warranted further investigation in a randomized cardiovascular outcome trial with the ultimate objective of ruling out a risk margin of 1.4. The LIGHT trial was adequately designed to achieve this objective. However, the Applicant has taken actions that have the potential to compromise the integrity of the LIGHT trial raising concerns about the ability to rely on data generated after the blind was broken to rule out the 1.4 risk margin. This in turn raises questions about the suitability of the LIGHT trial to achieve its ultimate objective in characterizing the CV risk of CONTRAVE. Due to these concerns one can postulate that the LIGHT trial is not being conducted to the highest of scientific standards. As such, we recommend a new cardiovascular outcome trial that is held to the highest of scientific standards be initiated with the objective of ruling out a relative CV risk of 1.4.

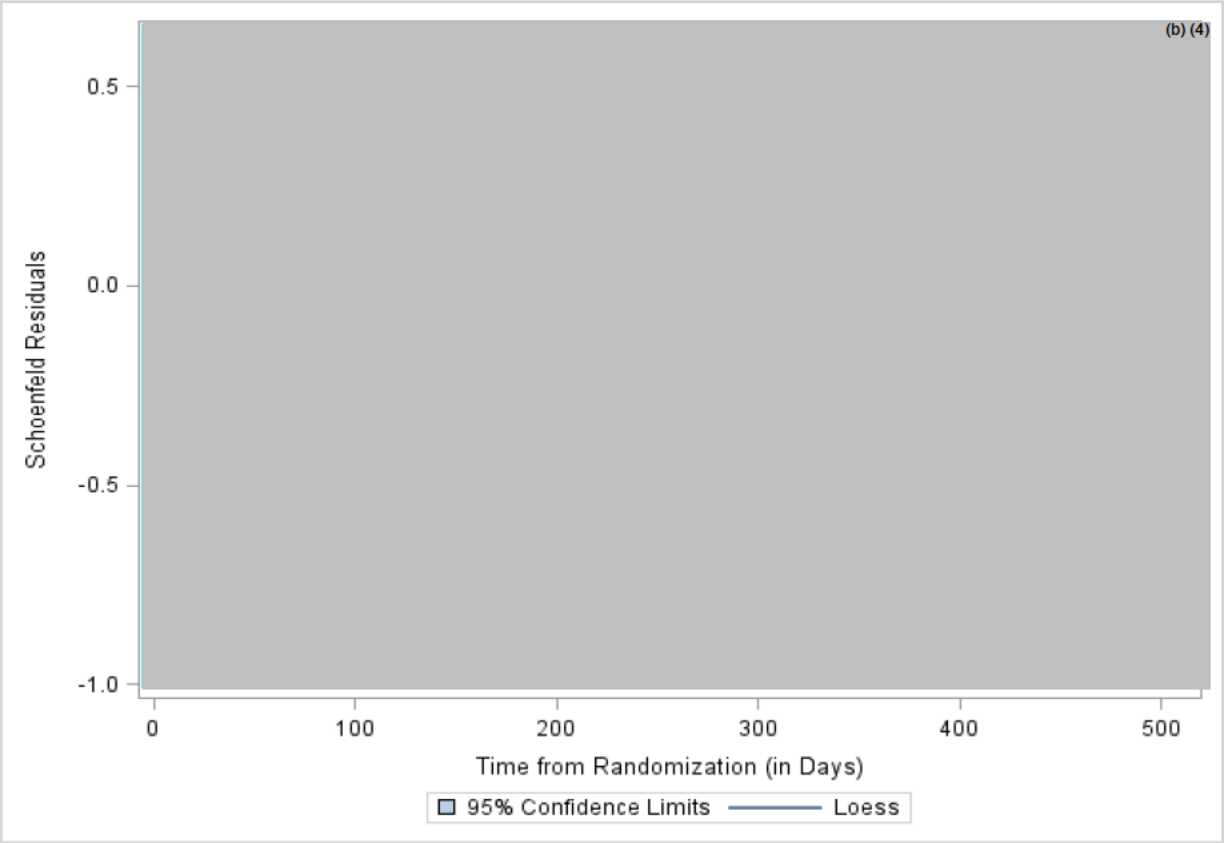
5.3 Labeling Recommendations

Regarding the CONTRAVE label, the recommendation is that none of the findings of the pre-approval interim analyses of the LIGHT trial be included in the label. Additionally, because the CV safety of CONTRAVE has not been confirmed with this interim data, the recommendation is that a limitation of use statement with respect to CV morbidity and mortality be included in the label. An example of such statement is "The effect of CONTRAVE on cardiovascular morbidity and mortality has not been established", which is consistent with labels for other currently approved weight loss products, namely, Belviq and Qsymia. If such a statement is included in the CONTRAVE label, the recommendation is that it be revisited after submission of the findings from a completed trial designed to assess the post-market risk margin of 1.4.

[28] Refer to Complete Response letter dated January 31, 2011.

APPENDIX 1 Assessment of Proportional Hazards Assumption

The plot below shows the Schoenfeld residuals and loess curve along with confidence bands used to evaluate the proportional hazards assumption for the Cox model analysis of MACE. A slope of zero is indicative that the proportional hazards assumption is not violated. Therefore, there are no major concerns about violation of this assumption from the interim analysis data.



Source: Created by the statistical reviewer using dataset "adtte.xpt"

APPENDIX 2 Additional Subjects included in Informative Censoring Sensitivity Analyses

The table below summarizes additional subjects included in the sensitivity analyses to investigate the potential for informative censoring in the on-study analyses. These subjects, (b) (4) are censored in the primary on-study MACE analysis, but are experiencing SAEs within 30 days of last contact date in the trial.

Subject ID	TRT	AE Term	AE Start Date	AE End Date	Last Contact Date
NB-CVOT-1047-10470034	(b) (4)	ARRHYTHMIA	21SEP2012	22OCT2012	15NOV2012
NB-CVOT-1148-11480125		CARDIOPULMONARY* ARREST	(b) (6)		

AE=adverse event, TRT=treatment (b) (4)
*Fatal event
Source: Created by the reviewer using datasets "adtte.xpt" and "adae.xpt"

The table below summarizes additional events included in the sensitivity analyses to investigate potential informative censoring in the on-treatment analyses. These subjects, (b) (4) CONTRAVE and (b) (4) placebo, are censored in the original on-treatment MACE analysis, but are experiencing SAEs within 30 days of last treatment date in the trial.

Subject ID	TRT	AE Term	AE Start Date	AE End Date	Treatment End Date
NB-CVOT-1002-10020005	(b) (4)	ANGINA PECTORIS	12SEP2012	13SEP2012	12OCT2012
NB-CVOT-1004-10040083		ATRIAL FLUTTER	21NOV2012	24NOV2012	29NOV2012
NB-CVOT-1012-10120071		NON-ST ELEVATION MYOCARDIAL INFARCTION	23FEB2013	27FEB2013	25FEB2013
NB-CVOT-1017-10170014		ACUTE CORONARY SYNDROME	27OCT2012	02NOV2012	03NOV2012
NB-CVOT-1022-10220012		ATRIAL FIBRILLATION	14AUG2013	.	29AUG2013
NB-CVOT-1027-10270055		MYOCARDIAL INFARCTION	07NOV2012	.	27JAN2013
NB-CVOT-		ATRIAL FIBRILLATION	01APR2013	04APR2013	01APR2013

Subject ID	TRT	AE Term	AE Start Date	AE End Date	Treatment End Date
1033-10330002	(b) (4)				
NB-CVOT-1035-10350100		ATRIAL FLUTTER	04FEB2013	15FEB2013	15FEB2013
NB-CVOT-1047-10470001		ATRIAL FIBRILLATION	13AUG2012	13AUG2012	17AUG2012
NB-CVOT-1051-10510033		CONGESTIVE HEART AILURE EXACERBATION	09SEP2012	05DEC2012	09SEP2012
NB-CVOT-1058-10580011		ATRIAL FIBRILLATION	11DEC2012	21DEC2012	14DEC2012
NB-CVOT-1062-10620072		EXACERBATION OF ATRIAL FIBRILLATION	05OCT2012	07OCT2012	23OCT2012
NB-CVOT-1064-10640057		VENTRICULAR TACHYCARDIA	21JUN2013	27JUN2013	20JUL2013
NB-CVOT-1174-11740041		SYMPTOMATIC CAROTID ARTERY STENOSIS	04JAN2013	17JAN2013	11JAN2013
NB-CVOT-1177-11770021		ATRIAL FIBRILLATION	11FEB2013	01MAR2013	11FEB2013
NB-CVOT-1197-11970033		OSSIBLE WORSENING OF NONISCHEMIC DILATED CARDIOMYOPATHY	27NOV2012	24JAN2013	31JAN2013
NB-CVOT-1256-12560035		CARDIAC CHEST PAIN	14JAN2013	24JAN2013	23JAN2013
NB-CVOT-1275-12750056		ATRIAL FIBRILLATION	11MAR2013	19MAR2013	11MAR2013
NB-CVOT-1295-12950005		RECURRENT UNSTABLE ANGINA	10DEC2012	14DEC2012	10DEC2012
NB-CVOT-1296-12960004		NONSUSTAINED VENTRICULAR TACHYCARDIA	05NOV2012	08APR2013	28NOV2012
NB-CVOT-1296-12960007		EXERTIONAL ANGINA	30NOV2012	01DEC2012	22DEC2012

AE=adverse event, TRT=treatment, NB=CONTRAVE, PBO=placebo
Source: Created by the reviewer using datasets “adtte.xpt” and “adae.xpt”

APPENDIX 3 Additional Subjects included in All-Cause Mortality Analysis

The table below summarizes 8 additional subjects (b) (4) who are included in the all-cause mortality analysis presented in this review, but are excluded from the Applicant's analysis in the study report.

Subject ID	TRT	AE TERM	Randomization Date	Death Date
NB-CVOT-1058-10580050	(b) (4)	DEATH	12SEP2012	(b) (6)
NB-CVOT-1069-10690119		DEATH	07NOV2012	
NB-CVOT-1074-10740077		CHRONIC SYSTOLIC CONGESTIVE HEART FAILURE	05OCT2012	
NB-CVOT-1078-10780044*		DEATH	24AUG2012	
NB-CVOT-1164-11640026		DEATH -UNKNOWN CAUSE	13NOV2012	
NB-CVOT-1249-12490011		MYOCARDIAL INFARCTION	20SEP2012	
NB-CVOT-1254-12540097		DEATH	06SEP2012	
NB-CVOT-1279-12790015		DEATH DUE TO MOTOR VEHICLE ACCIDNET	29OCT2012	

*Subject NB-CVOT-1078-10780044 was reported to have died but death date was not recorded. For this subject last contact date of (b) (6) was imputed as death date in reviewer's all-cause mortality analysis.
NB=CONTRAVE, PBO=placebo
Source: Created by the statistical reviewer using datasets "adtte.xpt", "adae.xpt", and "adsl.xpt"

APPENDIX 4 Projections for On-treatment Events at Trial Completion

Given the high discontinuation from treatment rates observed with the interim data, see Section 3.4, these projections are made to estimate the number of on-treatment events expected at completion of the LIGHT trial in 2.5 years, provided the trial is not terminated early. The table below summarizes these projections, based on the following assumptions and the fact that the trial is no longer enrolling additional subjects:

- Annual background on-treatment event rate of 1.10 per 100 person-years of exposure based on observed rates from interim data.
- Scenario #1: No additional subjects discontinue treatment, that is, 1709 CONTRAVE and 1201 placebo subjects continue on randomized treatment for remainder of trial.
- Scenario #2: Annual rate of treatment discontinuation, 60% for CONTRAVE and 70% placebo, for the four years of the trial.
- Scenario #3: Retention rates following year 1 are based on rates projected by the Applicant shown in Table 5, that is, 33% CONTRAVE and 25% placebo retained after year 2, 30% CONTRAVE and 22% PBO retained after year 3, and 28% CONTRAVE and 21% PBO retained after year 4.

	% subjects retained*	Total PYE*	Total Events*
Scenario #1	38% NB, 27% PBO	12180	134
Scenario #2	3% NB, 1% PBO	7500	83
Scenario #3	28% NB, 21% PBO	9400	104

NB=CONTRAVE, PBO=placebo, PYE=person-years of exposure, HR=hazard ratio
*Results expected after four years
Source: Created by the statistical reviewer

Note that the LIGHT trial is planned to observe 378 MACE for testing the 1.4 post-approval risk margin. Assuming a background event rate of 0.015, a minimum of 25200 person-years is needed to observe the planned MACE events. Note that based on the table of projected on-treatment events above, it is possible that the majority of the person-years, should the trial continue to completion, will be after the subjects have discontinued study medication. Therefore, this may introduce issues when interpreting the results at trial completion.

APPENDIX 5 List of Individuals with Access to Unblinded Interim Data

The tables below comprise the list of individuals provided by the Applicant, who, as of April 16, 2014, who were granted access to unblinded interim data of the ongoing LIGHT trials. In this list, "knowledge of threshold" refers to knowledge that the pre-approval threshold has been met and "full access" refers to access to unblinded summary and subject data; refer to Section 3.1 for detailed definitions of these terms. This list excludes potential (b) (4) with whom the Applicant has discussed unblinded data.

NAME	AFFILIATION	FUNCTION	INFORMATION ACCESSED
(b) (4)			

2 Page(s) has been Withheld in Full as b4 (CCI/TS) immediately following this page

SIGNATURE/DISTRIBTUION LIST

Primary Statistical Reviewer: Janelle K. Charles
Date: May 12, 2014

Concurring Reviewers: Mat Soukup, Team Leader DBVII
Aloka Chakravarty, Division Director DBVII

cc:
Curtis Rosebraugh, Director, Office of Drug Evaluation II
Jean-Marc Guettier, Director, DMEP
Eric Colman, Deputy Director, DMEP
Jennifer Pippins, Deputy Director for Safety (Acting), DMEP
James Smith, Medical Team Leader, DMEP
Eileen Craig, Medical Officer, DMEP
Patricia Madara, Project Manager, DMEP
Mehreen Hai, Project Manager (Safety), DMEP
Lisa LaVange, Director, Office of Biostatistics
Aloka Chakravarty, Director, DBVII
Mat Soukup, Statistics Team Leader, DBVII
Janelle Charles, Statistical Reviewer, DBVII
Lillian Patrician, Office of Biostatistics

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

/s/

JANELLE K CHARLES
05/12/2014

MATTHEW J SOUKUP
05/12/2014
Concur with review

ALOKA G CHAKRAVARTY
05/12/2014

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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	200-063 / SN 000
Drug Name:	Contrave (Naltrexone Hydrochloride and Bupropion Hydrochloride, Naltrexone HCl and Bupropion HCl, NB)
Indication(s):	Treatment of obesity and weight management, including weight loss and maintenance of weight loss
Applicant:	Orexigen Therapeutics Inc
Date(s):	Submitted March 31, 2010
Review Priority:	Standard
Biometrics Division:	Division of Biometrics VII
Statistical Reviewer:	Xiao Ding, Ph.D.
Concurring Reviewers:	Mat Soukup, Ph.D. Aloka Chakravarty, Ph.D.
Medical Division:	Division of Metabolism and Endocrinology Products (DMEP)
Clinical Team:	Medical Officer: Eileen Craig, M.D. Medical Team Leader: Eric Colman, M. D.
Project Manager:	Meghna Jairath (DMEP)

Keywords: NDA review, clinical studies, safety, vital sign

Table of Contents

1 Executive Summary 5
1.1 Conclusions and Recommendations 5
1.2 Brief Overview of Clinical Studies 6
1.3 Statistical Issues and Findings 6
2 Introduction 8
2.1 Product Description 8
2.2 Regulatory History 8
2.3 Clinical Trial Overview 8
2.4 Data Sources 9
3 Statistical Evaluation 9
3.1 Evaluation of Safety 9
3.1.1 Study Designs 10
3.1.2 Statistical Methodologies 11
3.1.2.1 Methods of Imputing Missing 11
3.1.2.2 Comparing treatment group difference 12
3.1.3 Subject Disposition, Demographic and Baseline Characteristics 12
3.1.4 Populations 15
3.1.5 Endpoints 16
3.1.5.1 Treatment-Emergent Adverse Event: Hypertension 16
3.1.5.2 Vital Signs 16
3.1.6 Results and Conclusions 17
3.1.6.1 Treatment-Emergent Hypertension 17
3.1.6.2 Vital Signs 19
3.1.6.2.1 Change of Vital Signs from Baseline to Week 56 19
3.1.6.2.2 Change of Vital Signs over Time 21
3.1.6.2.3 Change of Vital Signs by Treatment Response Group 26
4 Findings in Special/Subgroup Populations 31
4.1 Gender, Race, and Age 31
4.2 Other Special/Subgroup Populations 31
4.2.1 Subjects with Stable Weight 31
5 Summary and Conclusions 34
5.1 Statistical Issues and Collective Evidence 34
5.2 Conclusions and Recommendations 35
Signatures/Distribution List 37

Table of Tables

Table 1: Baseline Demographics by Study and Treatment Group (ITT Population) 13
Table 2: Study Discontinuation by Treatment Group (ITT Population)..... 15
Table 3: Treatment distribution in the Safety Population and the Completers Population 16
Table 4: Incidence of Treatment-Emergent Hypertension at Week 56 by Treatment Group (Safety Population) 17
Table 5: Time-to-Onset and Duration of Treatment-Emergent Hypertension (Safety Population, Double-Blinded Treatment Phase) 18
Table 6: Change of Pulse Rate from Baseline to Week 56 by Study (Safety Population, LOCF) 20
Table 7: Change of Systolic Blood Pressure (mm Hg) from Baseline to Week 56 by Study (Safety Population, LOCF)..... 20
Table 8: Change of Diastolic Blood Pressure (mm Hg) from Baseline to Week 56 by Study (Safety Population, LOCF)..... 21
Table 9: Change of Pulse Rate (bpm) from Baseline for Subjects with Measurement by Visit (Safety Population, Available Data at Each Visit with No Imputation) 21
Table 10: Change of Systolic Blood Pressure (mm Hg) from Baseline for Subjects with Measurement by Visit (Safety Population, Available Data at Each Visit with No Imputation)..... 23
Table 11: Change of Diastolic Blood Pressure (mm Hg) from Baseline for Subjects with Measurement by Visit (Safety Population, Available Data at Each Visit with No Imputation)..... 25
Table 12: Change of Vital Signs from Baseline to Week 56 by Treatment Response Group (Completers Population)..... 27

Table of Figures

Figure 1: Percentage of Subjects Who Discontinued from Study Drug for Any Reason by Study and Treatment Group (ITT Population)..... 14
Figure 2: Time to Onset of Treatment-Emergent Hypertension (Safety Population, Double-Blind Treatment Phase) 19
Figure 3: Mean Change of Pulse Rate over Time by Treatment Group (Safety Population, LOCF) 22
Figure 4: Mean Change of Pulse Rate over Time by Treatment Group (Completers Population) 22
Figure 5: Mean Change of Systolic Blood Pressure over Time by Treatment Group (Safety Population, LOCF) 24
Figure 6: Mean Change of Systolic Blood Pressure over Time by Treatment Group (Completers Population) 24
Figure 7: Mean Change of Diastolic Blood Pressure over Time by Treatment Group (Safety Population, LOCF) 25
Figure 8: Mean Change of Diastolic Blood Pressure over Time by Treatment Group (Completers Population) 26
Figure 9: Mean Change of Weight over Time by Treatment Response Group (Completers Population) 27
Figure 10: Mean Change of Pulse Rate by Treatment Response Group (Completers Population) 28
Figure 11: Mean Change of Systolic Blood Pressure by Treatment Response Group (Completers Population) 29
Figure 12: Mean Change of Diastolic Blood Pressure by Treatment Response Group (Completers Population) 30
Figure 13: Mean Change of Pressure-Rate Product by Treatment Response Group (Completers Population) 30
Figure 14: Mean Change of Pulse Rate by Treatment Group (Stable Weight Subpopulation) 31
Figure 15: Mean Change of Systolic Blood Pressure by Treatment Group (Stable Weight Subpopulation) 32
Figure 16: Mean Change of Diastolic Blood Pressure by Treatment Group (Stable Weight Subpopulation) 33
Figure 17: Mean Change of Pressure-Rate Product by Treatment Group (Stable Weight Subpopulation) 33

1 Executive Summary

This statistical review and evaluation was performed in response to a consultation from the Division of Metabolism and Endocrinology Products (DMEP) for New Drug Application (NDA) 200-063/000 (received March 31, 2010) for Naltrexone Hydrochloride and Bupropion Hydrochloride (NB) tablets. The proposed indication for NB is weight management, including weight loss and maintenance of weight loss. This statistical review assesses vital sign related safety parameters (pulse rate, systolic blood pressure, and diastolic blood pressure) in addition to treatment-emergent hypertension in the phase 3 clinical development program (Study NB-301, Study NB-302, Study NB-303, and Study NB-304) of NB tablets.

1.1 Conclusions and Recommendations

Based on the pooled analysis of the four phase 3 randomized placebo-controlled clinical studies NB-301, NB-302, NB-303, and NB-304, the NB32 treatment regimen (Naltrexone 32 mg Sustained Release/Bupropion 360 mg Sustained Release) was found to be associated with an increased risk of developing treatment-emergent hypertension as compared to the placebo regimen. The incidence of treatment-emergent hypertension was statistically significantly higher at the nominal α=0.05 level in the NB32 group than in the placebo group (*p=0.02*). Furthermore, the onset time of treatment-emergent hypertension was also statistically significantly earlier in the NB32 group than in the placebo group (*p=0.004*).

Compared to placebo, the NB32 treatment regimen was demonstrated to have less of a beneficial effect on vital signs, measured as change of pulse rate, systolic and diastolic blood pressure, from baseline to week 56. At 56 weeks, the mean change of pulse rate from baseline was statistically significantly higher in the NB32 group than in the placebo group, in three out of the four phase 3 studies separately, as well as in the pooled analysis of all four studies (0.91 bpm versus -0.20 bpm, with *p<0.0001*). The mean change of systolic blood pressure from baseline to 56 weeks was statistically significantly higher in the NB32 group than in the placebo group, in two out of the four phase 3 studies separately, as well as in the pooled analysis of all four studies (-0.22 mm Hg versus -1.62 mm Hg, with *p<0.0001*). Similarly, the mean change of diastolic blood pressure from baseline to 56 weeks was statistically significantly higher in the NB32 group than in the placebo group, in two out of the four phase 3 studies separately, as well as in the pooled analysis of all four studies (-0.76 mm Hg versus -1.35 mm Hg, with *p=0.1*).

Several exploratory analyses were done to explore the change of vital signs for treatment response groups (subjects who lost at least 5% of baseline body weight by the end of week 56 were identified as treatment responder). In the exploratory analyses, the placebo responders appeared to have the most beneficial change in vital signs (i.e. a reduction in the vital signs from baseline to week 56), while the NB32 non-responders tended to have the least beneficial change in vital signs.

1.2 Brief Overview of Clinical Studies

In this NDA application, the applicant submitted four phase 3 studies, Study NB-301, Study NB-302, Study NB-303, and Study NB-304, in support of the safety and efficacy of NB for the indication of weight management.

Both Study NB-301 and Study NB-303 were 56 week, multicenter, randomized, double blinded, placebo controlled, parallel-group studies to compare the safety and efficacy of NB and placebo in obese subjects. Study NB-302 was a 56-week, multicenter, randomized, double blinded, placebo controlled, parallel-group study to compare the safety and efficacy of NB and placebo in obese subjects *participating in a behavior modification program*. Study NB-304 was a 56-week, multicenter, randomized, double blinded, placebo controlled, parallel-group study to compare the safety and efficacy of NB and placebo in obese subjects *with type 2 diabetes mellitus*.

In the pooled data base of Studies NB-301, NB-302, NB-303, and NB-304, a total of 3,088 subjects were randomized to receive NB, while a total of 1,448 subjects were randomized to receive placebo. Details of the four phase 3 studies are provided in Section 3.1.1.

1.3 Statistical Issues and Findings

As pre-specified in the applicant's Statistical Analysis Plan (SAP), phase 3 data from Studies NB-301, NB-302, NB-303, and NB-304 were pooled to evaluate the risk of treatment-emergent hypertension and the changes of vital signs. The comparison between the NB32 group and the placebo group is performed using the protocol-defined Cochran-Mantel-Haenszel (CMH) test with study as stratification factor for the binary outcome, and the protocol-defined analysis of covariance (ANCOVA) model for continuous outcomes.

Compared to placebo, the NB32 regimen was found to be associated with an increased risk of developing treatment-emergent hypertension. The incidence of treatment-emergent hypertension was higher in the NB32 group than in the placebo group (6.0% versus 4.1%). The pooled relative risk ratio between NB32 and placebo was 1.41 with a 95% confidence interval of (1.05, 1.91). In addition to occurring at a higher frequency, treatment-emergent hypertension also appeared to occur earlier in the NB32 group than in the placebo group. The onset time of treatment-emergent hypertension was statistically significantly earlier for NB32 as compared to placebo, with p-value=0.0039 for stratified log-rank test. The stratified Cox proportional hazard ratio (HR) was 1.57 with 95% CI (1.15, 2.13). Detailed analysis results are provided in Section 3.1.6.1.

As compared to the placebo group, the mean change of pulse rate from baseline to week 56 was statistically significantly higher in the NB32 group, in Study NB-301, NB-302, and NB-303 separately, as well as in the pooled analysis of all four studies. Among the four phase 3 studies, the difference between NB32 and placebo in mean change of pulse rate ranged from 0.89 to 1.21 bpm indicating a higher pulse rate for NB32 treated

subjects. The pooled mean difference between NB32 and placebo was 1.11 bpm, with a 95% confidence interval of (0.62, 1.60). Detailed analysis results are provided in Section 3.1.6.2.1.

As compared to the placebo group, the mean change of systolic blood pressure from baseline to week 56 was statistically significantly higher in the NB32 group, in Study NB-301 and NB-302 separately, as well as in the pooled analysis of all four studies. The difference between NB32 and placebo in mean change of systolic blood pressure ranged from 0.67 to 2.63 mm Hg among the four studies indicating a higher systolic blood pressure in NB32 treated subjects. The pooled mean difference between NB32 and placebo was 1.41 mm Hg, with a 95% confidence interval of (0.76, 2.05). Detailed analysis results are provided in Section 3.1.6.2.1.

As compared to the placebo group, the mean change of diastolic blood pressure from baseline to week 56 was statistically significantly higher in the NB32 group, in Study NB-301 and NB-302 separately, as well as in the pooled analysis of all four studies. The difference between NB32 and placebo in mean change of diastolic blood pressure ranged from -0.01 to 1.32 mm Hg among the four studies indicating a higher diastolic blood pressure in NB32 treated subjects. The pooled mean difference between NB32 and placebo was 0.58 mm Hg, with a 95% confidence interval of (0.13, 1.03). Detailed analysis results are provided in Section 3.1.6.2.1.

In addition to the change from baseline to week 56, the changes of vital signs over time from baseline to each visit during the 56 weeks of study follow-up were also assessed in this review. In the pooled analysis of all four phase 3 studies, the changes of vital sign were shown to be higher in the NB32 group than in the placebo group throughout the 56 weeks of study follow up. The difference between NB32 and placebo appeared to be largest at early time points, and tended to decrease at later study visits. At week 56, the pooled difference between NB32 and placebo was 1.25 bpm in pulse rate, 1.30 mm Hg in systolic blood pressure, and 0.38 mm Hg in diastolic blood pressure, while the pooled difference at week 8 was 2.35 bpm, 2.31 mm Hg, and 1.98 mm Hg respectively. Detailed analysis results are provided in Section 3.1.6.2.2.

Based on several exploratory analyses which incorporate whether a subject responded to treatment or not (responders=subjects who lost at least 5% of baseline body weight by the end of week 56), the placebo responders appeared to have the most beneficial change in all vital signs from baseline to week 56. On average, the subjects in the placebo responder group had a reduction of 2.09 bpm in pulse rate, 4.85 mm Hg in systolic blood pressure, and 3.82 mm Hg in diastolic blood pressure. On the contrary, NB32 non-responders had the least beneficial change of vital signs from baseline to week 56. On average, the subjects in the NB32 non-responder group had an increase of 0.51 bpm and 0.80 mm Hg in pulse rate and systolic blood pressure respectively and only had a reduction of 0.06 mm Hg in diastolic blood pressure. More details of this assessment can be found in Section 3.1.6.2.3.

2 Introduction

2.1 Product Description

Naltrexone hydrochloride (naltrexone HCl) is an approved mu-opioid receptor antagonist indicated for the treatment of opiate and alcohol dependence, while bupropion hydrochloride (bupropion HCl) is an approved norepinephrine and dopamine reuptake inhibitor indicated for the treatment of major depression and nicotine dependence.

The applicant has developed a combination product of naltrexone HCl and bupropion HCl sustained-release tablets. The proposed indication of Naltrexone/Bupropion (NB) tablets is for treatment of obesity and weight management, including weight loss and maintenance of weight loss, in conjunction with lifestyle modification. NB tablets are proposed for patients with an initial body mass index ≥30 kg/m2 or ≥27 kg/m2 with one or more risk factors (e.g., diabetes, dyslipidemia, or hypertension).

2.2 Regulatory History

Bupropion HCl was first approved by the FDA for the treatment of depression in 1985, smoking cessation in 1997, and seasonal affective disorder in 2006. Bupropion HCl is currently marketed as three formulations: an immediate-release tablet, a sustained-release tablet, and an extended-release tablet.

Naltrexone HCl was approved by the FDA in 1984. Naltrexone immediate-release tablet formulation is currently approved for the treatment of opioid addiction (1984) and alcohol dependence (1995). In 2006, an extended-release injectable suspension formulation of naltrexone HCl was approved for the treatment of alcohol dependence.

2.3 Clinical Trial Overview

The applicant submitted the results of four phase 3, randomized, controlled clinical trials (NB-301, NB-302, NB-303, and NB-304) in support of the safety and efficacy of NB tablets for treatment of obesity and weight management indication.

Study NB-301 was entitled “A Multicenter, Randomized, Double Blind, Placebo Controlled Study Comparing the Safety and Efficacy of Two Doses of Naltrexone Sustained Release (SR)/Bupropion Sustained Release (SR) and Placebo in Obese Subjects”.

Study NB-302 was entitled “A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Comparing the Safety and Efficacy of Naltrexone Sustained Release (SR)/Bupropion Sustained Release (SR) and Placebo in Subjects with Obesity Participating in a Behavior Modification Program”.

Study NB-303 was entitled "A Multicenter, Randomized, Double Blind, Placebo Controlled Study Comparing the Safety and Efficacy of Naltrexone Sustained Release (SR)/Bupropion Sustained Release (SR) and Placebo in Obese Subjects".

Study NB-304 was entitled "A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Comparing the Safety and Efficacy of Naltrexone 32 mg Sustained Release/Bupropion 360 mg Sustained Release and Placebo in Obese Subjects with Type 2 Diabetes Mellitus".

Among these four phase 3 studies, Studies NB-301 and NB-303 were conducted in obese subjects receiving customary diet and behavioral counseling. Study NB-302 was conducted in obese subjects undergoing intensive lifestyle modification counseling. Study NB-304 was conducted in obese subjects with type 2 diabetes.

Study NB-301 investigated two daily doses of NB (naltrexone 16 mg/bupropion 360 mg [NB16] and naltrexone 32 mg/bupropion 360 mg [NB32]). Study NB-302 assessed the efficacy and safety of NB32 in a population of obese subjects undergoing an intensive behavioral modification program that included prescribed diet, exercise, and counseling. In Study NB-303, subjects who did not experience or maintain at least 5% weight loss between Weeks 28-44 of NB32 therapy were re-randomized to continue on NB32 or increase their daily dose to naltrexone 48 mg/bupropion 360 mg (NB48) to determine if the dose increase resulted in a therapeutic benefit.

2.4 Data Sources

The applicant submitted electronic documents and datasets for Studies NB-301, NB-302, NB-303 and NB-304. The following files available within the CDER Electronic Document Room (EDR) were utilized in this review.

\\Cdsesub1\evsprod\NDA200063\0000\m5\datasets\iss
\\Cdsesub1\evsprod\NDA200063\0000\m5\datasets\nb-301
\\Cdsesub1\evsprod\NDA200063\0000\m5\datasets\nb-302
\\Cdsesub1\evsprod\NDA200063\0000\m5\datasets\nb-303
\\Cdsesub1\evsprod\NDA200063\0000\m5\datasets\nb-304

3 Statistical Evaluation

This review is focused on specific safety parameters, specifically evaluation of vital signs including blood pressure and heart rate. For a complete statistical evaluation of efficacy results, please refer to the review authored by Dr. Janice Derr.

3.1 Evaluation of Safety

The review of safety comprises data from Studies NB-301, NB-302, NB-303 and NB-304. Based upon interactions with the clinical review team, the following review of safety consists of a focused evaluation of vital signs and treatment-emergent

hypertension. All comparative analyses are between the randomized treatment groups, NB and Placebo.

3.1.1 Study Designs

Study NB-301 was a phase 3 multicenter, randomized, double blind placebo-controlled study involving 3 treatment groups treated for approximately 56 weeks followed by a 2-week drug discontinuation period. NB-301 was conducted at 34 sites in the United States, in male and female subjects between 18 to 65 years of age with either uncomplicated obesity or with obesity/overweight with controlled hypertension and/or dyslipidemia. A total of 1742 subjects were randomized in a 1:1:1 fashion to one of the three treatment groups: placebo, NB16, and NB32. The randomization was stratified by study center to achieve the desired overall allocation of treatments balanced across the participating centers. Study NB-301 was comprised of four periods: a screening period of up to 4 weeks (at least 2 visits), a titration period of 4 weeks (1 visit); a study drug maintenance period of 52 weeks (14 visits); and a drug discontinuation period of 2 weeks (1 visit) for a total of 58 weeks of study duration. Subjects were to be seen every 4 weeks from baseline to Week 56, and at Week 58 following the 2-week drug discontinuation period. At the Week 56 visit (completion of maintenance period), subjects in each active treatment group were re-randomized in a double-blind 1:1 fashion to undergo either tapered withdrawal or sudden withdrawal of study drug. Placebo subjects were not re-randomized at this point but continued to receive blinded placebo treatment to maintain study blind during the discontinuation treatment period. Study NB-301 was powered to provide 99% or greater power to detect a statistically significant difference between placebo and the NB arms for the percent change from baseline to Week 56 of total body weight.

Study NB-302 was a phase 3 multicenter, randomized, double-blind, placebo-controlled study involving 2 treatment groups treated for 56 weeks. NB-302 was conducted at 9 sites in the United States, in male and female subjects between 18 to 65 years of age with either uncomplicated obesity or with obesity/overweight with controlled hypertension and/or dyslipidemia who were concurrently participating in an intense behavior modification program that included dietary instructions, twenty-eight 90-minute group sessions, and prescribed exercise. A total of 793 subjects were randomized in a 3:1 fashion to NB32 and placebo. The randomization was stratified by study center to achieve the desired overall allocation of treatments balanced across the participating centers. Study NB-302 was powered to provide 99% or greater power to detect a statistically significant difference between placebo and NB arm for the percent change from baseline to Week 56 of total body weight.

Study NB-303 was a phase 3 multicenter, randomized, double-blind, placebo-controlled study involving 2 treatment groups treated for 56 weeks. NB-303 was conducted at 36 sites in the United States, in male and female subjects between 18 to 65 years of age with uncomplicated obesity or with obesity/overweight with controlled hypertension and/or dyslipidemia. A total of 1,496 subjects were randomized in a 2:1 fashion to NB32 and placebo. The randomization was stratified by study center to achieve the desired overall allocation of treatments balanced across the participating centers. Subjects randomized to

NB32 arm were randomized within study center in a 1:1 ratio to two alternative titration schedules of naltrexone SR (fast vs. slow). Beginning at Week 28 through Week 44, NB32-treated subjects who failed to achieve or maintain at least 5% body weight loss from baseline were re-randomized in a 1:1 fashion to continue NB32 or begin treatment with a higher dose of NB48. Study NB-303 was powered to provide 99% or greater power to detect a statistically significant difference between placebo and NB arm for the percent change from baseline to Week 56 of total body weight.

Study NB-304 was a phase 3 multicenter, randomized, double-blind, placebo-controlled study involving 2 treatment groups treated for 56 weeks. NB-303 was conducted at 53 sites in the United States, in obese/overweight subjects between 18 and 70 years of age with type 2 diabetes (not on injectable diabetes medications or inhaled insulin). A total of 505 subjects were randomized in a 2:1 fashion to NB32 and placebo. Randomization was centrally stratified by baseline HbA1c (≤8% or >8%) and pharmacotherapy (with or without sulfonylurea) to achieve the desired overall allocation of treatments across participating centers. Study NB-304 was powered to provide 99% or greater power to detect a statistically significant difference between placebo and NB arm for the percent change from baseline to Week 56 of total body weight.

3.1.2 Statistical Methodologies

In the following sections, statistical methods and tabulations are presented for the evaluation of safety only.

3.1.2.1 Methods of Imputing Missing

Subjects who did not experience resolution of a specific treatment-emergent adverse event were right-censored at the date of last confirmed dose date + 7 days. Therefore the missing value of the stop date was imputed as the last confirmed dose date +7 days.

Reviewer's comment: *The applicant imputed the missing stop date of a certain treatment-emergent adverse event as the last confirmed dose date. Because of the definition of treatment-emergent adverse event, this review imputes the missing stop date as the last confirmed dose date + 7 days.*

For subjects with at least one post baseline measurement, the last observation carried forward (LOCF) method was used to impute the missing value of vital signs.

Reviewer's comment: *Although the LOCF method to impute the missing values was pre-specified in the protocol, this method can lead to biased point estimates and variances. This is especially problematic when the study discontinuation rate is high. Therefore sensitivity analyses based on the completers population were conducted to assess the robustness of the primary results based on the safety population and the LOCF imputation.*

3.1.2.2 Comparing treatment group difference

For treatment-emergent hypertension, the comparison of proportions between the NB32 group and the placebo group is performed using the protocol-defined Cochran-Mantel-Haenszel (CMH) test with study as the stratification factor. The two-sided 95% confidence interval for between-group relative risk ratio was conducted by a stratified Mantel-Haenszel approach based on pooled phase 3 data from Studies NB-301, NB-302, NB-303, and NB-304. For each individual study, the two-sided 95% confidence interval was conducted using the normal approximation to the binomial distribution.

Kaplan-Meier method was used to demonstrate the cumulative probability of initial occurrence of treatment-emergent hypertension. The comparison of the curves between treatment groups was assessed by the protocol-defined stratified log-rank test with study as the stratification factor.

For vital signs, the change from baseline value was compared between treatment groups using the protocol-defined analysis of covariance (ANCOVA) model containing terms for treatment, study, with the appropriate baseline measurement as a covariate (least squares [LS] Means and Type III).

In the exploratory analysis of vital signs by treatment responder (defined as subjects who lost at least 5% of baseline body weight by the end of week 56), no formal inference is conducted between treatment responder groups. Rather graphical methods are used to assess trends over time for treatment groups based upon subjects efficacy responder status.

3.1.3 Subject Disposition, Demographic and Baseline Characteristics

As shown in Table 1, in the intent-to-treat population (ITT population), defined as all randomized subjects, baseline demographics and characteristics were similar among the treatment groups. All subjects in Studies NB-301, NB-302 and NB-303 were between the ages of 18 and 66 years, while subjects with type 2 diabetes mellitus in Study NB-304 were between the ages of 20 and 72 years. More than 85% of subjects in the studies were female, except for Study NB-304 where the distribution of males and females was more balanced (56% female). In the ITT population, approximately 77% of subjects were Caucasian and about 18% were Black or African American, while there were about 10% subjects with Hispanic or Latino ethnicity. In Studies NB-301, NB-302 and NB-303, the mean weight was about 100 kg while the subjects in the NB-304 study had a slightly higher weight (105 kg). A higher percentage of subjects with type 2 diabetes mellitus had hypertension (62%) and dyslipidemia (84%) compared to the other studies (NB-301, NB-302, and NB-303) where 20% of subjects had hypertension and 50% had dyslipidemia. In the latter studies, about 25% of subjects had impaired fasting glucose at the baseline visit.

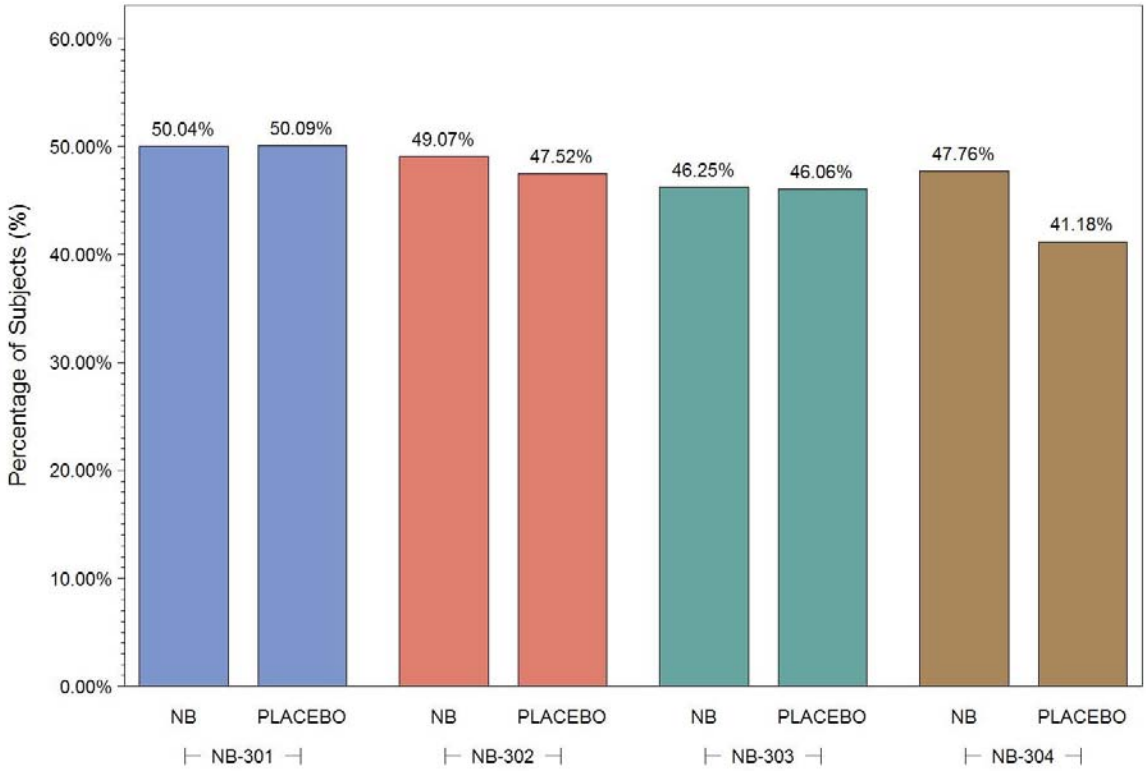
Table 1: Baseline Demographics by Study and Treatment Group (ITT Population)

	Study NB-301 (N=1742)		**Study NB-302 (N=793)**		**Study NB-303 (N=1496)**		**Study NB-304 (N=505)**	
	Placebo n=581	**NB*** n=1161	**Placebo** n=202	**NB** n=591	**Placebo** n=495	**NB** n=1001	**Placebo** n=170	**NB** n=335
Age (years) Mean (SD)	43.7 (11.1)	44.4 (11.2)	45.6 (11.4)	45.9 (10.4)	44.4 (11.4)	44.3 (11.2)	53.5 (9.8)	54.0 (9.1)
Weight (kg) Mean (SD)	99.4 (14.3)	99.6 (15.3)	101.9 (15.0)	100.2 (15.4)	99.2 (15.9)	100.3 (16.5)	105.1 (17.0)	104.2 (18.9)
BMI (kg/m^2) Mean (SD)	36.2 (4.0)	36.2 (4.3)	37.0 (4.2)	36.3 (4.2)	36.1 (4.3)	36.2 (4.4)	36.4 (4.5)	36.4 (4.7)
Gender, (%)								
Female	85.4%	84.9%	91.6%	89.3%	84.9%	84.6%	52.9%	58.2%
Male	14.6%	15.1%	8.4%	10.7%	12.1%	15.4%	47.1%	41.8%
Race, (%)								
White	75.7%	74.7%	73.8%	68.5%	83.6%	83.4%	82.3%	77.9%
Black	18.9%	19.6%	21.8%	24.5%	13.6%	13.3%	10.6%	18.8%
Others	5.3%	5.7%	4.4%	7.0%	1.8%	3.3%	7.1%	3.3%
Ethnicity, (%) Hispanic/Latino	14.6%	12.4%	9.9%	9.6%	9.1%	6.9%	11.8%	11.3%
Hypertension (%)	19.6%	21.4%	18.3%	14.6%	21.4%	21.2%	60.6%	63.3%
Dyslipdemia, (%)	49.6%	49.2%	40.1%	45.7%	53.1%	55.9%	85.3%	83.6%
Impaired Fast Glucose, (%)	24.1%	25.8%	24.3%	21.7%	25.7%	28.0%	type 2 diabetes	type 2 diabetes

* In Study NB-301, the NB arm included both the NB16 group and the NB32 group
Source: Created by reviewer.

Among the 4,536 randomized subjects in the ITT population from Studies NB-301, NB-302, NB-303, and NB-304, approximately 48% of them (2179/4536) discontinued study prior to week 56. As presented in Figure 1, the percentage of subjects who discontinued treatment for any reason was comparable overall between the placebo and the NB treatment groups. Across all four studies, the incidence of treatment discontinuation tended to be higher in the NB treatment groups than in the placebo groups. In Studies NB-301, NB-302 and NB-303, the discontinuation rate was close between the treatment groups. In Study NB-304, the incidence of treatment discontinuation was noticeably higher in the NB treatment group than in the placebo group (47.8% versus 41.2%), but the difference is not statistically significant (*p=0.18* Fisher's exact test).

Figure 1: Percentage of Subjects Who Discontinued from Study Drug for Any Reason by Study and Treatment Group (ITT Population)



* In Study NB-301, the NB arm included both the NB16 group and the NB32 group
Source: Created by reviewer.

As presented in Table 2, the most common reasons reported for study discontinuation were adverse event, lost to follow-up, and withdrawal of consent, which accounted for 19.8%, 9.9%, and 9.9% of subjects in the ITT population respectively. In each of the studies NB-301, NB-302, NB-303, and NB-304, the proportion of study discontinuation due to adverse event was higher in the NB group than in the placebo group. On average, 23.4% of the subjects in the NB group discontinued study due to adverse events, which was statistically significantly higher than the placebo group (12.1%) with p-value<0.0001 (CMH test stratified by studies). On the contrary, the proportion of study discontinuation due to lack of efficacy was lower in the NB group than in the placebo group in each of the four studies. On average, 1.7% of the subjects in the NB group discontinued study due to lack of efficacy, which was statistically significantly lower than the placebo group (5.9%) with p-value<0.0001 (CMH test stratified by studies).

Table 2: Study Discontinuation by Treatment Group (ITT Population)

	Study NB-301 (N=1742)		Study NB-302 (N=793)		Study NB-303 (N=1496)		Study NB-304 (N=505)	
	Placebo n=581	**NB*** n=1161	**Placebo** n=202	**NB** n=591	**Placebo** n=495	**NB** n=1001	**Placebo** n=170	**NB** n=335
Discontinued before week 56	291 (50.1)	581 (50.0)	96 (47.5)	106 (52.5)	228 (46.1)	463 (46.3)	70 (41.2)	160 (47.8)
Adverse Event	56 (9.6)	234 (20.2)	25 (12.4)	150 (25.4)	68 (13.7)	241 (24.1)	26 (15.3)	98 (29.3)
Lost to Follow-up	73 (12.6)	147 (12.7)	22 (10.9)	27 (4.6)	53 (10.7)	90 (9.0)	15 (8.8)	22 (6.6)
Withdrawal of Consent	90 (15.5)	123 (10.6)	24 (11.9)	43 (7.3)	56 (11.3)	75 (7.5)	15 (8.8)	21 (6.3)
Lack of Efficacy	40 (6.9)	24 (2.1)	6 (3.0)	3 (0.5)	33 (6.7)	19 (1.9)	6 (3.5)	5 (1.5)
Non-compliance with treatment	15 (2.6)	25 (2.2)	5 (2.5)	13 (2.2)	5 (1.0)	10 (1.0)	3 (1.8)	8 (2.4)
Protocol Violation	10 (1.7)	14 (1.2)	0 (0)	7 (1.2)	10 (2.0)	21 (2.1)	5 (2.9)	3 (0.9)
Other	7 (1.2)	14 (1.2)	2 (1.0)	6 (1.0)	3 (0.6)	7 (0.7)	0 (0)	3 (0.9)

* In Study NB-301, the NB arm included both the NB16 group and the NB32 group
Source: Created by reviewer.

3.1.4 Populations

The analyses for all safety outcomes will be primarily based on the safety population. The safety population includes all randomized subjects who received at least one tablet of study treatment and have at least one investigator contact or assessment at any time after the start of study treatment, regardless of whether or not they discontinued the study.

Subjects who completed a full 56 weeks of study drug treatment (completers population) were also evaluated. For studies NB-301, NB-303, and NB-304, this analysis set included all randomized subjects with a baseline measurement, a postbaseline body weight measurement, and who completed 56 weeks of treatment. For Study NB-302, the completer analysis set included all randomized subjects who had a baseline measurement and a post-baseline measurement at Week 56 while on study drug (i.e. active treatment).

As shown in Table 3, among the 4,536 randomized subjects in the ITT population from Studies NB-301, NB-302, NB-303, and NB-304, 4,481 (98.8%) subjects were included in the safety population and 2,357 (52.0%) subjects were included in the completers population.

Table 3: Treatment distribution in the Safety Population and the Completers Population

	Placebo	NB16	NB32	Total NB
Number of Subjects Randomized (ITT Population)	N=1448	N=578	N=2510	N=3088
Safety Population	1430 (98.8%)	569 (98.4%)	2482 (98.9%)	3051 (98.8%)
Completers Population	763 (52.7%)	284 (49.1%)	1310 (52.2%)	1594 (51.6%)

Source: Created by reviewer.

3.1.5 Endpoints

3.1.5.1 Treatment-Emergent Adverse Event: Hypertension

The analysis of adverse event in this review primarily focused on Treatment-Emergent Adverse Event (TEAS), defined as adverse events that occurred or worsened on or after the date of first dose until 7 days after the last confirmed dose.

The definition of hypertension was based on all the components of 'hypertension' in the Standardized MedDRA[1] Queries (SMQ), including all the following preferred terms: Hypertension, blood pressure increased, ECG signs of ventricular hypertrophy, labile hypertension, blood pressure diastolic increased, blood pressure systolic increased, and cardiovascular disorder.

The time to onset of a specific treatment-emergent adverse event was calculated (in days) as the difference between the start date of that adverse event and the date of the first dose of study treatment +1 day. The duration of a specific treatment-emergent adverse event was defined as the difference between the stop date and the start date of that AE + 1 day.

3.1.5.2 Vital Signs

Vital signs, including pulse rate, systolic blood pressure, and diastolic blood pressure, were measured in the sitting position at every study visit (Baseline, Week 4, Week 8, Week 12, Week 16, Week 20, Week 24, Week 28, Week 32, Week 36, Week 40, Week 44, Week 48, Week 52, Week 56). In Studies NB-301, NB-302, NB-303 and NB-304, the value used for analyses was an average of three blood pressure and pulse readings obtained after the subject had been sitting for at least 5 minutes to minimize the effects of variability associated with measurement of blood pressure and pulse rate.

[1] Adverse events were recorded by the applicant using MedDRA dictionary version 12.0.

3.1.6 Results and Conclusions

3.1.6.1 Treatment-Emergent Hypertension

Only considering subjects randomized to receive NB32 or placebo in the safety population, the incidence of treatment-emergent hypertension (defined using the MedDRA SMQ – see Section 3.1.5.1) at week 56 is presented in Table 4. The incidence of treatment-emergent hypertension was greater in the NB32 group compared to the placebo group (6.0% vs. 4.1%). Pooling all four phase 3 studies NB-301, NB-302, NB-303 and NB-304 together, the stratified Mantel-Haenszel relative risk ratio between the NB32 group and the placebo group was 1.41 with a 95% CI of (1.05, 1.91), with p-value=002 (CMH test stratified by studies).

Table 4: Incidence of Treatment-Emergent Hypertension at Week 56 by Treatment Group (Safety Population)

	Study NB-301		**Study NB-301**		**Study NB-301**		**Study NB-301**	
	Placebo	NB32	Placebo	NB32	Placebo	NB32	Placebo	NB32
N of subjects	569	573	200	584	492	992	169	333
n of events	21	29	11	39	15	39	11	41
Incidence	3.7%	5.1%	5.5%	6.7%	3.1%	3.9%	6.5%	12.3%
Relative Risk (95% CI)	1.37 (0.79, 2.38)		1.21 (0.63, 2.32)		1.29 (0.72, 2.32)		1.89 (1.00, 3.58)	
Mantel-Haenszel Relative Risk (95% CI)	1.41 (1.05, 1.91) CMH p-value=0.02							

Source: Created by reviewer.

The time-to-onset and duration of treatment-emergent hypertension is summarized in Table 5. Among the subjects with treatment-emergent hypertension reported, 13.5% of subjects in the NB32 group experienced the initial onset within the first 4 weeks, compared to only 3.5% of subjects in the placebo group. Median time to initial onset was 17.5 and 26.5 weeks for NB32 and placebo, respectively, and the median duration was 8.5 weeks in the NB32 group compared to 5 weeks for the placebo group.

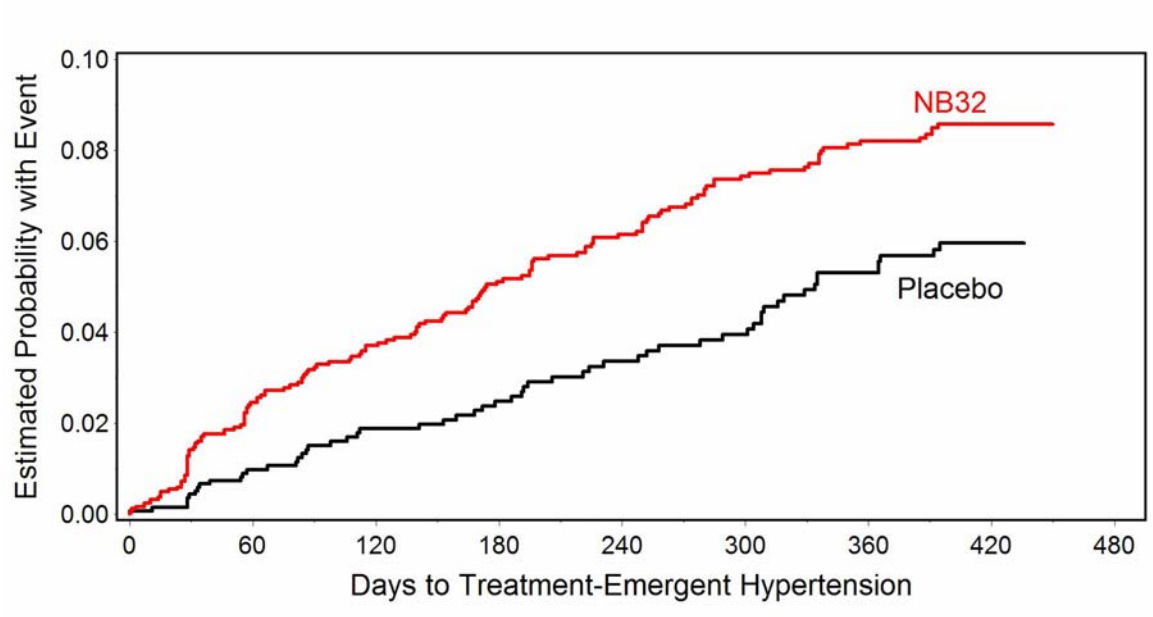
Table 5: Time-to-Onset and Duration of Treatment-Emergent Hypertension (Safety Population, Double-Blinded Treatment Phase)

	Placebo (N=1430)	**NB32 (N=2482)**
Subjects with Treatment-Emergent Hypertension n (%)	58 (4.1%)	148 (6.0%)
Time-to-onset (weeks)		
>0 to 4 weeks	2 (3.5%)	20 (13.5%)
>4 to 8 weeks	10 (17.2%)	23 (15.5%)
>8 to 12 weeks	4 (6.9%)	17 (11.5%)
>12 to 16 weeks	6 (10.3%)	10 (6.8%)
>16 to 20 weeks	1 (1.7%)	9 (6.1%)
>20 to 24 weeks	3 (5.2%)	9 (6.1%)
>24 to 28 weeks	6 (10.3%)	13 (8.8%)
>28 to 32 weeks	3 (5.2%)	8 (5.4%)
>32 to 36 weeks	3 (5.2%)	8 (5.4%)
>36 to 40 weeks	2 (3.5%)	9 (6.1%)
>40 to 44 weeks	4 (6.9%)	7 (4.7%)
>44 to 48 weeks	7 (12.1%)	3 (2.0%)
>48 to 52 weeks	2 (3.5%)	7 (4.7%)
>52 to 56 weeks	3 (5.2%)	4 (2.7%)
>56 weeks	2 (3.5%)	1 (0.7%)
Mean (SD)	27.0 (17.2)	21.0 (16.0)
Median	26.5	17.5
Duration (weeks)		
Mean (SD)	12.2 (13.9)	13.8 (14.6)
Median	5.0	8.5

Source: Created by reviewer.

Based on the Kaplan Meier method using safety information from the four phase 3 trials, the cumulative probability of developing treatment-emergent hypertension is shown in Figure 2. In the NB32 group, treatment-emergent hypertension occurred earlier and more often than in the placebo group. Based on the stratified log-rank test stratified by study, the onset time of treatment-emergent hypertension was statistically significantly different between the NB32 group and the placebo group (p-value=0.004). The stratified Cox proportional hazard ratio (HR) was 1.57 with 95% CI (1.15, 2.13).

Figure 2: Time to Onset of Treatment-Emergent Hypertension (Safety Population, Double-Blind Treatment Phase)



Source: Created by reviewer.

3.1.6.2 Vital Signs

3.1.6.2.1 Change of Vital Signs from Baseline to Week 56

Pulse Rate

For the subjects in the safety population who had at least one post baseline measurement of pulse rate, the results of the change in pulse rate from baseline to week 56 for the four phase 3 studies are presented in Table 6. In all four phase 3 studies, the mean pulse rate increased from baseline to week 56 in the NB32 group, but decreased in the placebo group. The difference between NB32 and placebo ranged from 0.89 to 1.21 beat per minute (bpm). Pooling all four studies together, the difference in pulse rate between NB32 and placebo was 1.11 bpm with a 95% CI of (0.62, 1.60).

Table 6: Change of Pulse Rate from Baseline to Week 56 by Study (Safety Population, LOCF)

	Treatment	N	n with post baseline measurement	LS Mean (bpm)	Difference of LS Mean from Placebo (95% CI)	p-value * vs. Placebo
Study NB-301	Placebo	569	518	-0.29		
	NB32	573	491	0.93	1.21 (0.34, 2.08)	0.006
Study NB-302	Placebo	200	194	-0.17		
	NB32	584	509	0.87	1.03 (-0.17, 2.24)	0.09
Study NB-303	Placebo	492	464	-0.36		
	NB32	992	874	0.77	1.13 (0.35, 1.91)	0.005
Study NB-304	Placebo	169	161	-0.03		
	NB-32	333	293	0.87	0.89 (-0.49, 2.28)	0.21
Pooled Phase 3 Studies	**Placebo**	1430	1337	-0.20		
	NB32	2482	2167	0.91	1.11 (0.62, 1.60)	<0.0001

* Test of Type III sum of squares from ANCOVA model
Source: Created by reviewer.

Systolic Blood Pressure

For the subjects in the safety population who had at least one post baseline measurement of systolic blood pressure, the results of the change in systolic blood pressure from baseline to week 56 for the four phase 3 studies are presented in Table 7. In studies NB-301, NB-302, NB-303 and NB-304, the systolic blood pressure decreased more in the placebo group than in the NB32 group. The difference between NB32 and placebo in systolic blood pressure ranged from 0.67 to 2.63 mm Hg among the studies. Pooling all four studies together, the difference in systolic blood pressure between NB32 and placebo was 1.41 mm Hg with a 95% CI of (0.76, 2.05).

Table 7: Change of Systolic Blood Pressure (mm Hg) from Baseline to Week 56 by Study (Safety Population, LOCF)

	Treatment	N	n with post baseline measurement	LS Mean (mm Hg)	Difference of LS Mean from Placebo (95% CI)	p-value * vs. Placebo
Study NB-301	Placebo	569	518	-2.15		
	NB32	573	491	-0.46	1.69 (0.59, 2.78)	0.003
Study NB-302	Placebo	200	194	-3.84		
	NB32	584	509	-1.21	2.63 (1.08, 4.18)	0.0009
Study NB-303	Placebo	492	464	-0.54		
	NB32	992	874	0.13	0.67 (-0.32, 1.67)	0.18
Study NB-304	Placebo	169	161	-1.47		
	NB-32	333	293	-0.18	1.29 (-0.86, 3.45)	0.24
Pooled Phase 3 Studies	**Placebo**	1430	1337	-0.22		
	NB32	2482	2167	-1.62	1.41 (0.76, 2.05)	<0.0001

* Test of Type III sum of squares from ANCOVA model
Source: Created by reviewer.

Diastolic Blood Pressure

For the subjects in the safety population who had at least one post baseline measurement of diastolic blood pressure, the results of the change in diastolic blood pressure from baseline to week 56 for the four phase 3 studies are presented in Table 8. In studies NB-301, NB-302, and NB-304, the diastolic blood pressure decreased more in the placebo group than in the NB32 group. In study NB-303, there was only a marginal change of diastolic blood pressure from baseline to week 56 in both the NB32 and placebo groups. The difference between NB32 and placebo in diastolic blood pressure ranged from -0.01 to 1.32 mm Hg among the studies. Pooling all four studies together, the difference in diastolic blood pressure change from baseline to week 56 between NB32 and placebo was 0.58 mm Hg with a 95% CI of (0.13, 1.03).

Table 8: Change of Diastolic Blood Pressure (mm Hg) from Baseline to Week 56 by Study (Safety Population, LOCF)

	Treatment	N	n with post baseline measurement	LS Mean (mm Hg)	Difference of LS Mean from Placebo (95% CI)	p-value * vs. Placebo
Study NB-301	Placebo	569	518	-1.42		
	NB32	573	491	-0.43	0.99 (0.21, 1.76)	0.01
Study NB-302	Placebo	200	194	-2.59		
	NB32	584	509	-1.27	1.32 (0.20, 2.44)	0.02
Study NB-303	Placebo	492	464	-0.05		
	NB32	992	874	-0.05	-0.01 (-0.72, 0.71)	0.99
Study NB-304	Placebo	169	161	-1.62		
	NB-32	333	293	-1.29	0.33 (-1.06, 1.71)	0.65
Pooled Phase 3 Studies	**Placebo**	1430	1337	-0.76		
	NB32	2482	2167	-1.35	0.58 (0.13, 1.03)	0.01

* Test of Type III sum of squares from ANCOVA model
Source: Created by reviewer.

3.1.6.2.2 Change of Vital Signs over Time

Pulse Rate

For the subjects in the safety population with available data at each visit, the changes in pulse rate from baseline to different visits (weeks 4, 8, 12, and 56) are presented in Table 9. The pooled difference in pulse rate between NB32 and placebo was smallest at week 56 (1.25 bpm) compared to week 8 or week 12 values (2.35 or 2.63 bpm). One explanation for this observation is that as weight decreases it is expected that pulse rate decreases.

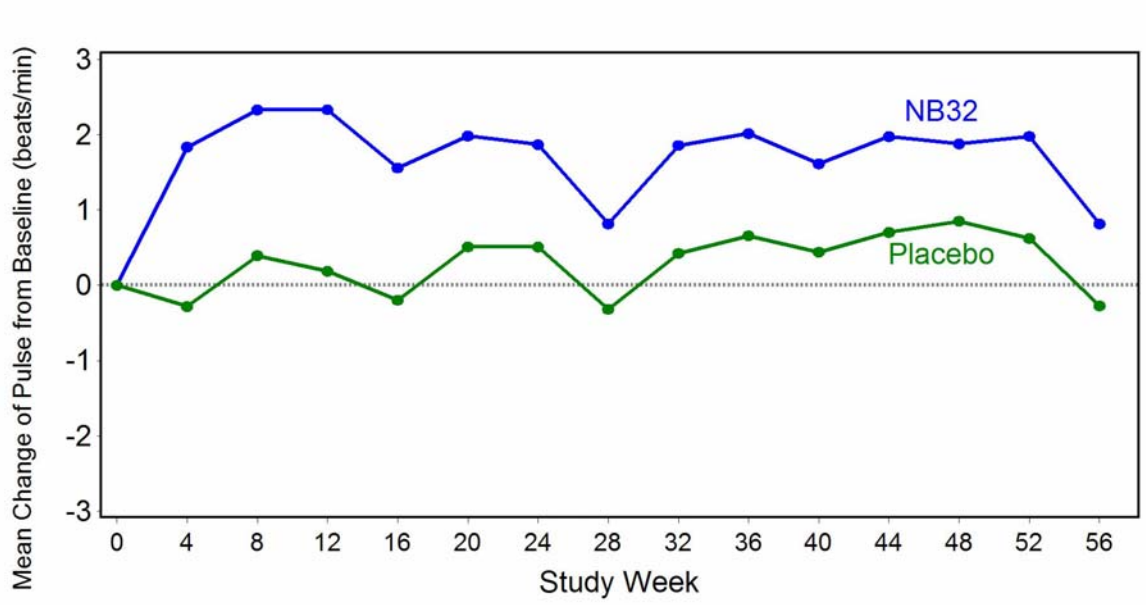
Table 9: Change of Pulse Rate (bpm) from Baseline for Subjects with Measurement by Visit (Safety Population, Available Data at Each Visit with No Imputation)

Visit	Treatment	n	LS Mean (bpm)	Difference of LS Mean from Placebo (95% CI)	p-value * (vs. Placebo)
Week 4	Placebo	1333	-0.16		
	NB32	2155	1.72	1.88 (1.25,2.51)	<0.0001
Week 8	Placebo	1208	0.14		
	NB32	1833	2.49	2.35 (1.67, 3.02)	<0.0001
Week 12	Placebo	1107	-0.07		
	NB32	1723	2.56	2.63 (1.97, 3.29)	<0.0001
Week 56	Placebo	751	-1.06		
	NB32	1312	0.19	1.25 (0.63, 1.87)	<0.0001

* Test of Type III sum of squares from ANCOVA model
Source: Created by reviewer.

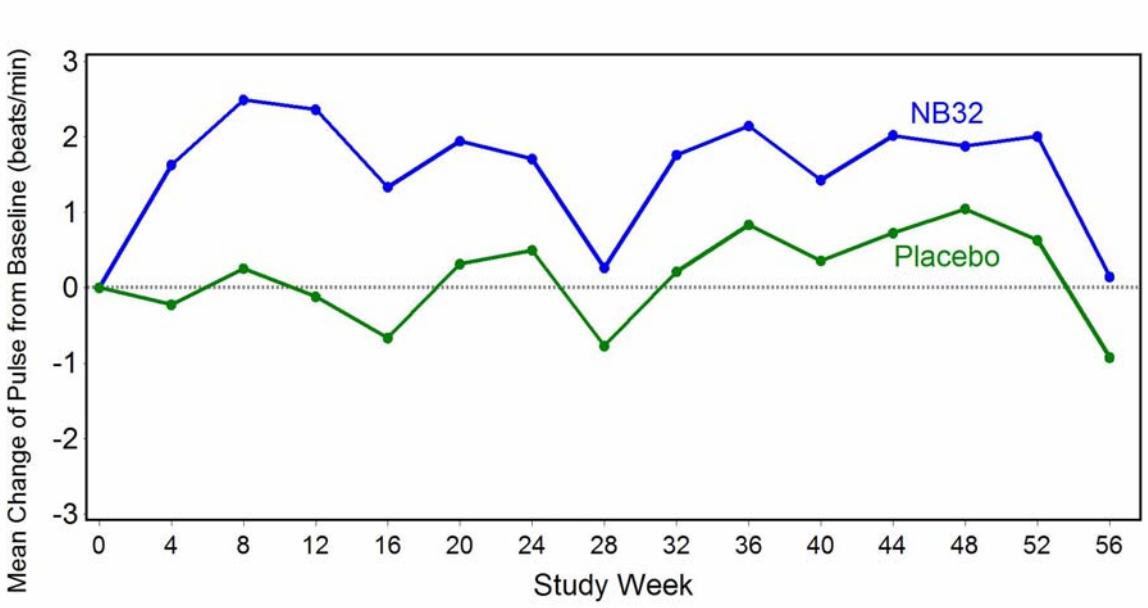
The mean change of pulse rate from baseline to each visit over time in the safety population is shown in Figure 3. The curve for NB32 was always above the curve for placebo. The difference between the two curves was largest at early time points, and tended to decrease at later study visits. In order to assess the impact of the LOCF imputation method on the results shown in Figure 3, the mean change of pulse rate from baseline to each visit over time in the completers population is shown in Figure 4. The pattern demonstrated in Figure 4 is similar to that shown in Figure 3.

Figure 3: Mean Change of Pulse Rate over Time by Treatment Group (Safety Population, LOCF)



Source: Created by reviewer.

Figure 4: Mean Change of Pulse Rate over Time by Treatment Group (Completers Population)



Source: Created by reviewer.

Systolic Blood Pressure

For the subjects in the safety population, the mean changes in systolic blood pressure from baseline to different visits (weeks 4, 8, 12, and 56) are presented in Table 10 for subjects with available data at each visit. The pooled difference in systolic blood pressure between NB32 and placebo was smallest at week 56 (1.30 mm Hg), as compared to week 8 or week 12 values (2.31 or 1.99 mm Hg). Note that mean systolic blood pressure at weeks 12 and 56 in the NB32 treatment group were below the baseline values. One explanation for this observation is that as weight decreases it is expected that systolic blood pressure decreases.

Table 10: Change of Systolic Blood Pressure (mm Hg) from Baseline for Subjects with Measurement by Visit (Safety Population, Available Data at Each Visit with No Imputation)

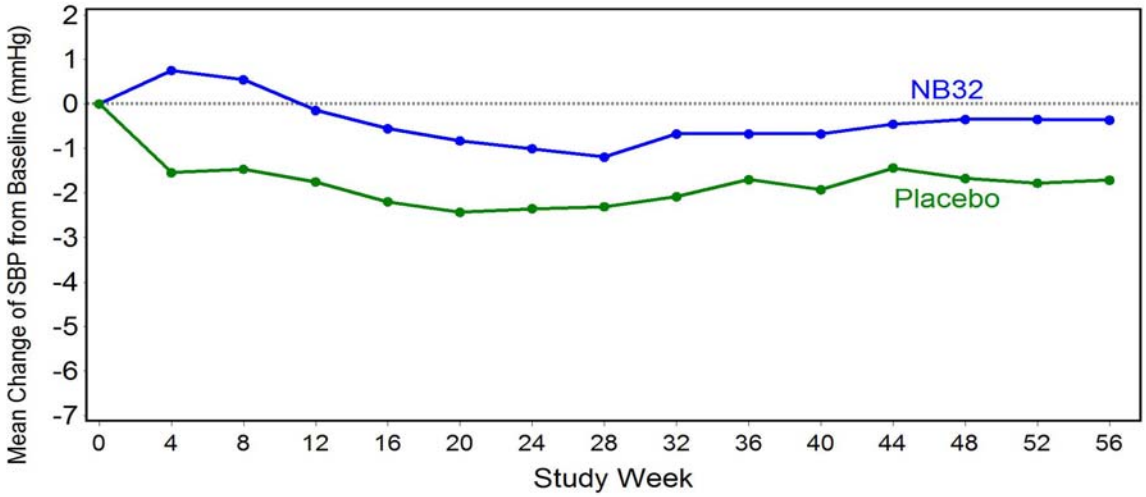
Visit	Treatment	n	LS Mean (mm Hg)	Difference of LS Mean from Placebo (95% CI)	p-value * (vs. Placebo)
Week 4	Placebo	1333	-1.33		
	NB32	2155	1.01	2.35 (1.60, 3.09)	<0.0001
Week 8	Placebo	1208	-1.69		
	NB32	1833	0.63	2.31 (1.55, 3.08)	<0.0001
Week 12	Placebo	1107	-2.10		
	NB32	1723	-0.11	1.99 (1.22, 2.77)	<0.0001
Week 56	Placebo	751	-1.98		
	NB32	1312	-0.68	1.30 (0.63, 1.87)	0.003

* Test of Type III sum of squares from ANCOVA model
Source: Created by reviewer.

The mean change of systolic blood pressure from baseline to each visit over time is shown in Figure 5. The curve for NB32 was always above the curve for placebo. The

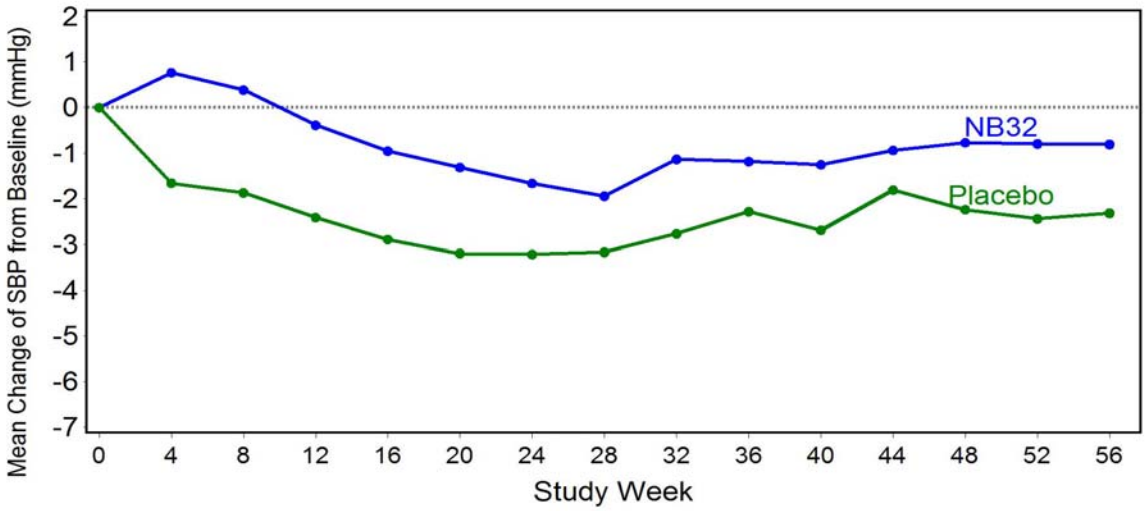
difference between the two curves was largest at early time points, and tended to decrease at later study visits. At visits beyond week 8, the mean change in systolic blood pressure was below zero indicating a decrease in mean systolic blood pressure from the baseline value. In order to assess the impact of the LOCF imputation method on the results shown in Figure 5, the mean change of systolic blood pressure from baseline to each visit over time in the completers population is shown in Figure 6. The pattern demonstrated in Figure 6 is similar to that shown in Figure 5.

Figure 5: Mean Change of Systolic Blood Pressure over Time by Treatment Group (Safety Population, LOCF)



Source: Created by reviewer.

Figure 6: Mean Change of Systolic Blood Pressure over Time by Treatment Group (Completers Population)



Source: Created by reviewer.

Diastolic Blood Pressure

For the subjects in the safety population with available data at each visit, the changes in mean diastolic blood pressure from baseline to different visits (weeks 4, 8, 12, and 56) are presented in Table 11. The pooled mean difference in diastolic blood pressure between NB32 and placebo was smallest at week 56 (0.38 mm Hg) as compared to week 8 or week 12 values (1.98 or 1.62 mm Hg). Note that mean diastolic blood pressure at weeks 12 and 56 in the NB32 treatment group were below the baseline values. One explanation for this observation is that as weight decreases it is expected that diastolic blood pressure decreases.

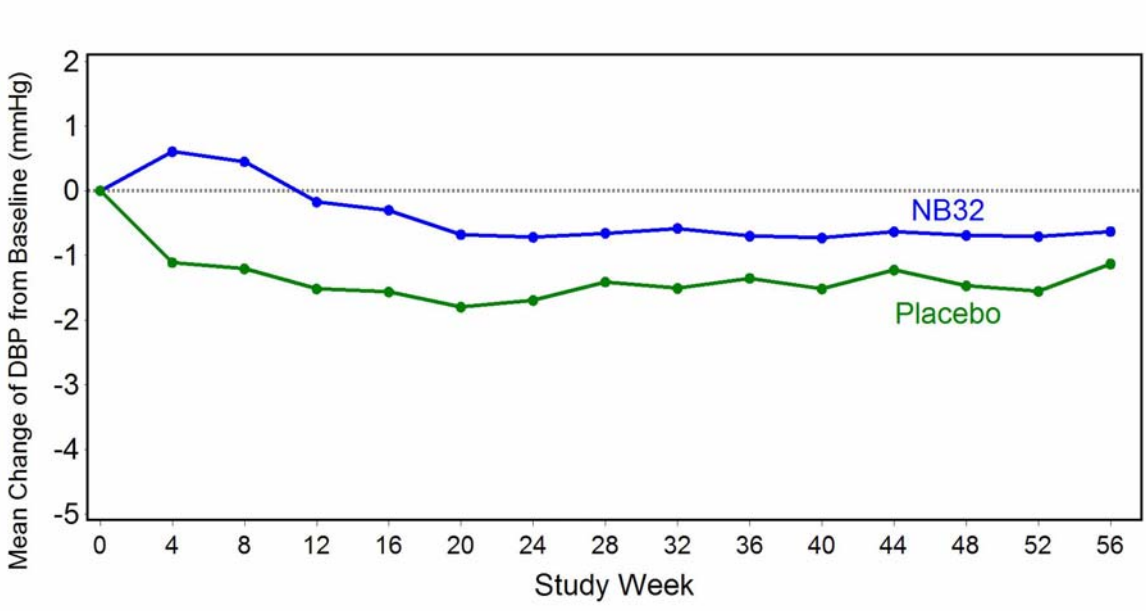
Table 11: Change of Diastolic Blood Pressure (mm Hg) from Baseline for Subjects with Measurement by Visit (Safety Population, Available Data at Each Visit with No Imputation)

Visit	Treatment	n	LS Mean (mm Hg)	Difference of LS Mean from Placebo (95% CI)	p-value * (vs. Placebo)
Week 4	Placebo	1333	-1.28		
	NB32	2155	0.63	1.90 (1.36, 2.45)	<0.0001
Week 8	Placebo	1208	-1.77		
	NB32	1833	0.20	1.98 (1.42, 2.53)	<0.0001
Week 12	Placebo	1107	-2.12		
	NB32	1723	-0.51	1.62 (1.07, 2.17)	<0.0001
Week 56	Placebo	751	-1.70		
	NB32	1312	-1.32	0.38 (-0.21, 0.97)	0.003

* Test of Type III sum of squares from ANCOVA model
Source: Created by reviewer.

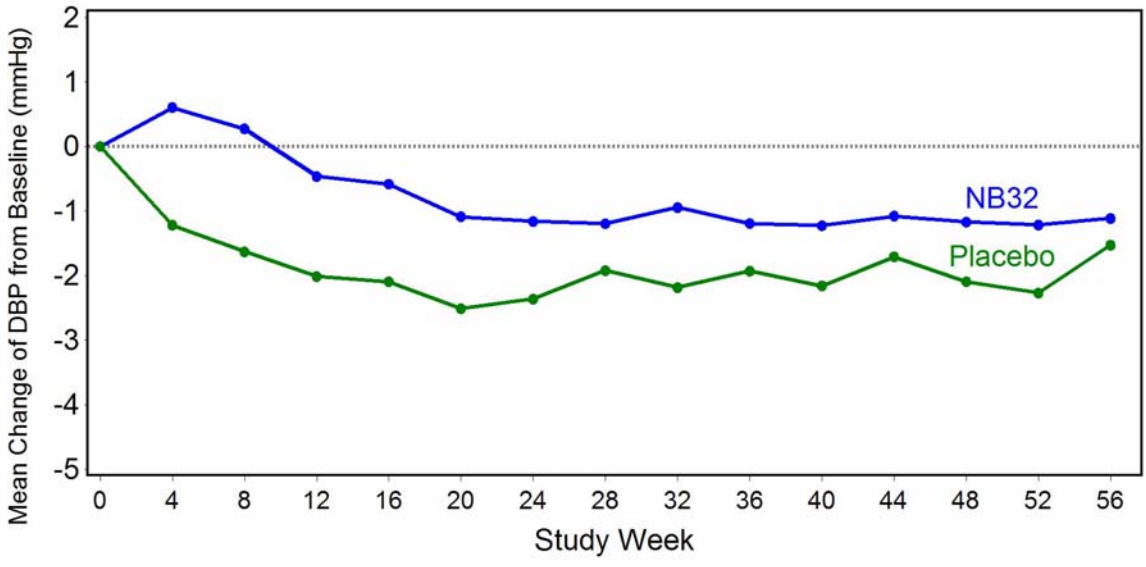
The mean change of diastolic blood pressure from baseline to each visit over time is shown in Figure 7. The curve for NB32 was always above the curve for placebo. The difference between the two curves was largest at early time points, and tended to decrease at later study visits. At visits beyond week 8, the mean change in diastolic blood pressure was below zero indicating a decrease in mean diastolic blood pressure from the baseline value. In order to assess the impact of the LOCF imputation method on the results shown in Figure 7, the mean change of diastolic blood pressure from baseline to each visit over time in the completers population is shown in Figure 8. The pattern demonstrated in Figure 8 is similar to that shown in Figure 7.

Figure 7: Mean Change of Diastolic Blood Pressure over Time by Treatment Group (Safety Population, LOCF)



Source: Created by reviewer.

Figure 8: Mean Change of Diastolic Blood Pressure over Time by Treatment Group (Completers Population)



Source: Created by reviewer.

3.1.6.2.3 Change of Vital Signs by Treatment Response Group

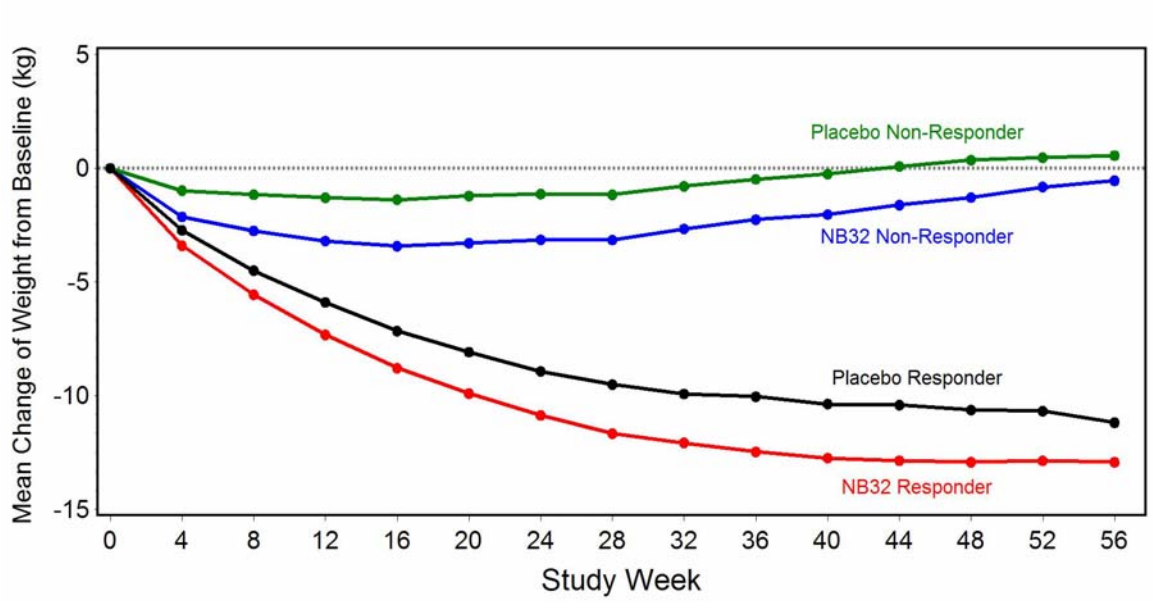
In this section, several exploratory analyses were done to explore the change of vital signs by treatment response group. According to the protocol, subjects who lost at least 5% of baseline body weight by the end of week 56 were considered as responders. In the completers population, 871 subjects randomized to NB32 and 212 subjects randomized to

placebo were identified as responders, while 439 subjects in the NB32 group and 551 subjects in the placebo group were identified as non-responders.

***Reviewer's comment**: Because the responders and non-responders were defined by post randomization variable, no formal statistical inference was done here. The comparison regarding treatment and the responder/non-responder subgroups should be interpreted with caution.*

For subjects in the completers population, the mean change of weight from baseline to each visit over time is shown in Figure 9. Over the 56 weeks treatment period, subjects in the NB32 responder group on average lost the most weight, followed by subjects in the placebo responder group. Subjects in the placebo non-responder group lost the least weight.

Figure 9: Mean Change of Weight over Time by Treatment Response Group (Completers Population)



Source: Created by reviewer.

For subjects in the completers population, the change of vital signs from baseline to week 56 was summarized in Table 12 for the NB32 responder group, the NB32 non-responder group, the placebo responder group, and the placebo non-responder group.

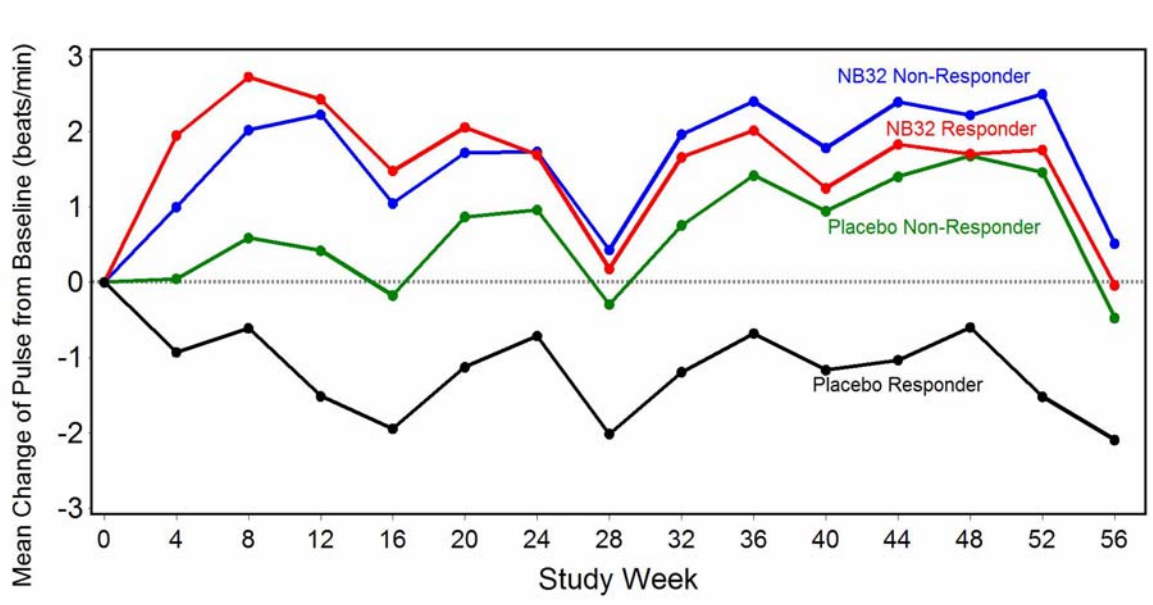
Table 12: Change of Vital Signs from Baseline to Week 56 by Treatment Response Group (Completers Population)

Mean (SD) Median (Range)	**Change of Pulse Rate (bpm)**	**Change of Systolic BP (mm Hg)**	**Change of Diastolic BP (mm Hg)**
NB32 Non-responder (N=439)	0.51 (7.97) 1 (-27, 31)	0.80 (10.44) 0 (-58, 39)	-0.06 (6.99) 0 (-28, 21)
Placebo Non-responder (N=551)	-0.47 (7.12) 0 (-26, 20)	-1.34 (9.88) -1 (-35, 33)	-0.65 (6.87) -1 (-24, 38)
NB32 Responder (N=871)	-0.04 (7.65) 0 (-31, 32)	-1.61 (9.74) -2 (-35, 32)	-1.65 (7.20) -2 (-30, 32)
Placebo Responder (N=212)	-2.09 (7.78) -2 (-23, 27)	-4.85 (11.24) -4 (-43, 43)	-3.82 (7.78) -3 (-29, 28)

The subjects in the placebo responder group had the most beneficial change for all vital signs from baseline to week 56. On average, they lost 2.09 bpm in pulse rate, 4.85 mm Hg in systolic blood pressure, and 3.82 mm Hg in diastolic blood pressure. On the contrary, the subjects in the NB32 non-responder group had the least beneficial change for all vital signs from baseline to week 56. On average, these NB32 non-responders gained 0.51 bpm and 0.80 mm Hg in pulse rate and systolic blood pressure respectively, and only lost 0.06 mm Hg in diastolic blood pressure.

For subjects in the completers population, the mean change of pulse rate from baseline to each visit over time is shown in Figure 10 by the NB32 and placebo response groups. Over the 56 weeks period, subjects in the placebo response group had the most beneficial mean changes in pulse rate, while the NB32 non-responders had the least beneficial mean changes.

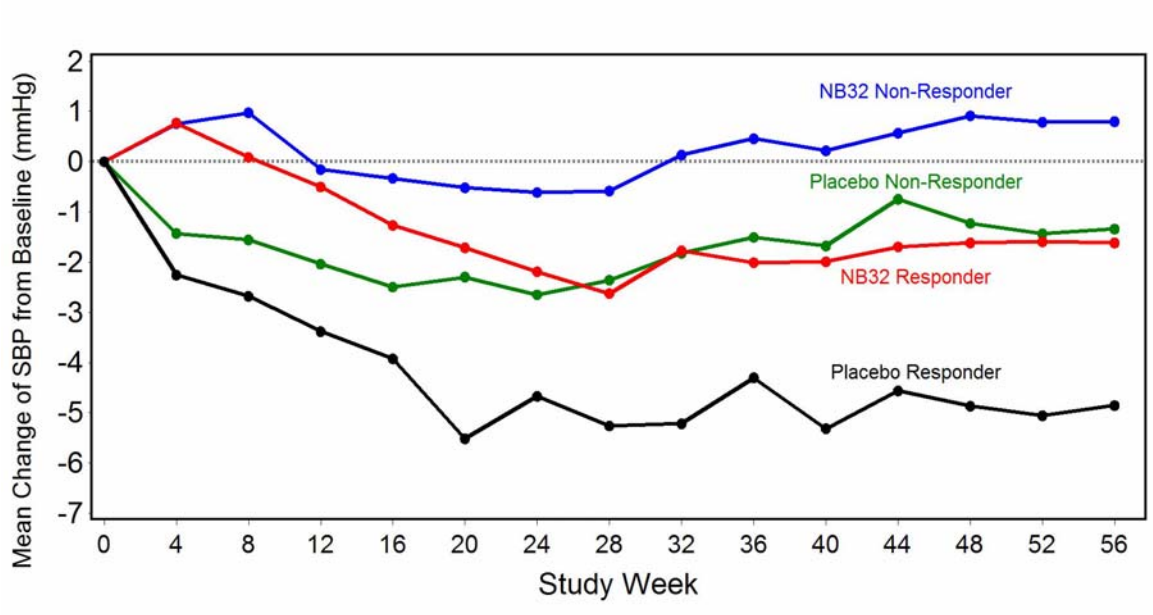
Figure 10: Mean Change of Pulse Rate by Treatment Response Group (Completers Population)



Source: Created by reviewer.

For subjects in the completers population, the mean change of systolic blood pressure from baseline to each visit over time is shown in Figure 11 by the NB32 and placebo response groups. Over the 56 weeks period, subjects in the placebo response group had the most beneficial mean changes in systolic blood pressure, while the NB32 non-responders had the least beneficial mean changes.

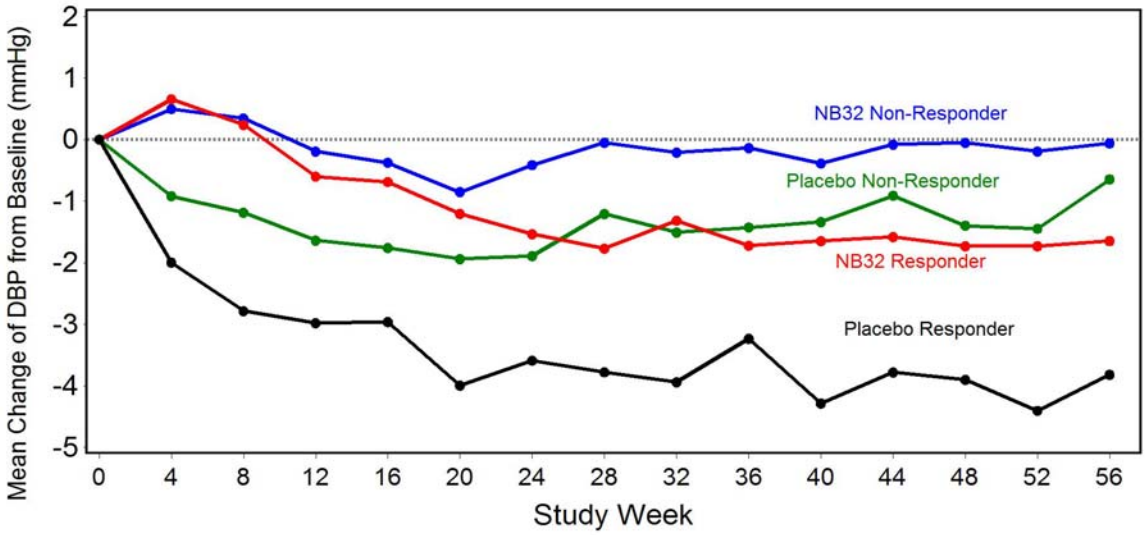
Figure 11: Mean Change of Systolic Blood Pressure by Treatment Response Group (Completers Population)



Source: Created by reviewer.

For subjects in the completers population, the mean change of diastolic blood pressure from baseline to each visit over time is shown in Figure 12 by the NB32 and placebo response groups. Over the 56 weeks period, subjects in the placebo response group had the most beneficial mean changes in diastolic blood pressure, while the NB32 non-responders had the least beneficial mean changes.

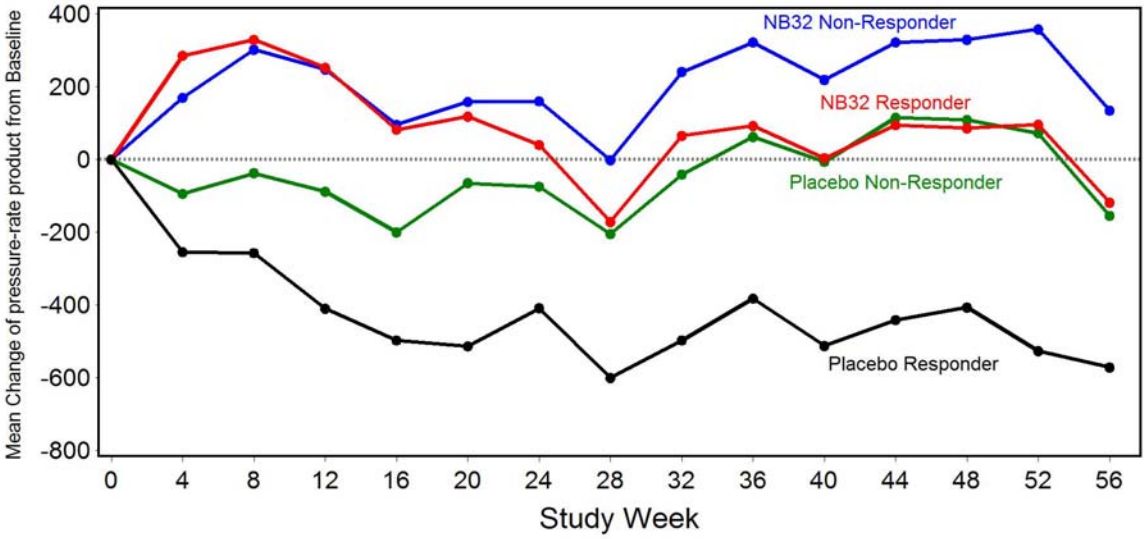
Figure 12: Mean Change of Diastolic Blood Pressure by Treatment Response Group (Completers Population)



Source: Created by reviewer.

Other than pulse rate, systolic and diastolic blood pressure, pressure-rate product (defined as the product of pulse rate and systolic blood pressure) is another vital sign parameter of clinical interest. For subjects in the completers population, the mean change of pressure-rate product from baseline to each visit over time is shown in Figure 13 by the NB32 and placebo response groups. Over the 56 weeks treatment period, subjects in the placebo response group had the most beneficial mean changes in pressure-rate product, while the NB32 non-responders had the least beneficial mean changes.

Figure 13: Mean Change of Pressure-Rate Product by Treatment Response Group (Completers Population)



Source: Created by reviewer.

4 Findings in Special/Subgroup Populations

4.1 Gender, Race, and Age

No subgroup analyses by gender, race, or age were assessed in this review.

4.2 Other Special/Subgroup Populations

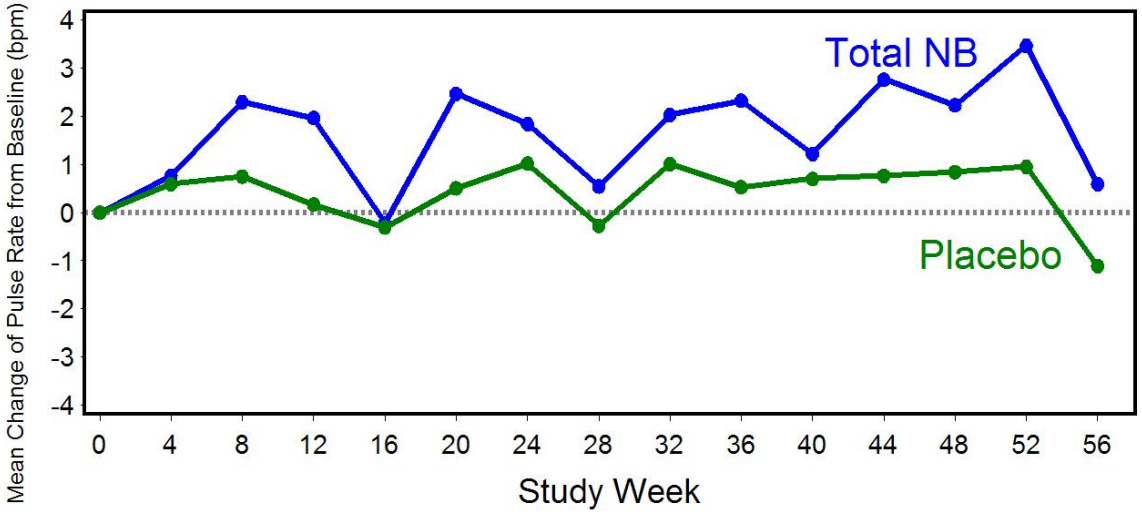
4.2.1 Subjects with Stable Weight

In order to exclude the impact of weight change on vital signs, it is of clinical interest to explore the change of vital signs for subjects with stable weight. Subjects whose weight was within ± 2% of the baseline value throughout the first 6 months were considered as having stable weight based on discussions with the clinical review team. In the Completers Population, a total of 192 subjects were identified as subjects with stable weight, with 64 in the total NB group (19 randomized to NB16 and 45 to NB32) and 128 in the placebo group.

***Reviewer's comment**: Because of the small number of subjects with stable weight in the NB32 group,, the total NB group including both the NB32 and the NB 16 groups was studied. Also note that since the stable weight subpopulation was defined by a post randomization variable, no formal statistical inference was done here. The comparison between NB and placebo in this subpopulation should be interpreted with caution.*

For the subjects with stable weight, the mean change of pulse rate from baseline to each visit over time is shown in Figure 14. At each visit over the 56 weeks period, the subjects with stable weight on placebo consistently had the more beneficial mean changes in pulse rate than those with stable weight on NB treatment.

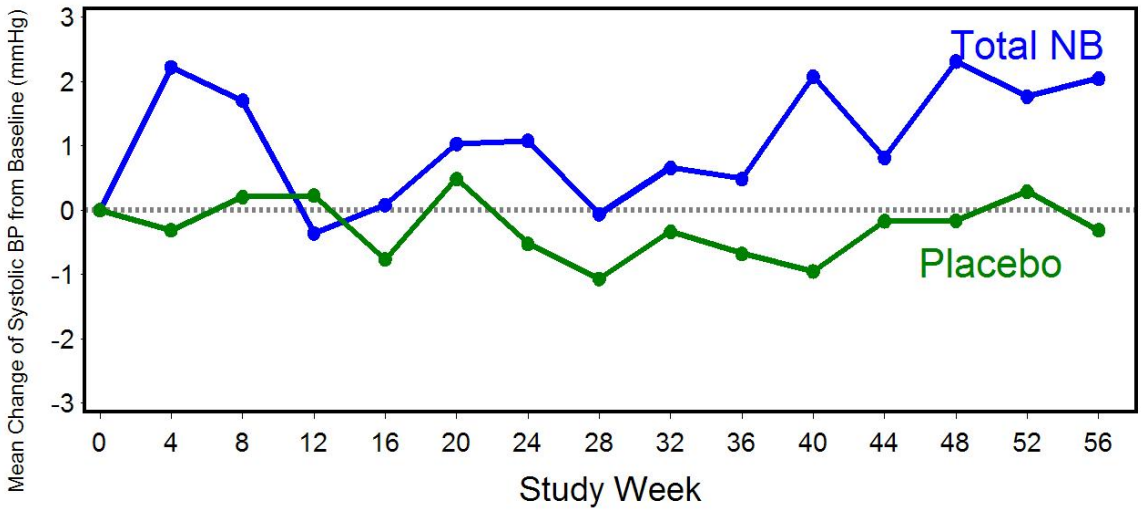
Figure 14: Mean Change of Pulse Rate by Treatment Group (Stable Weight Subpopulation)



Source: Created by reviewer.

Similarly, the mean changes of systolic blood pressure from baseline to each visit over time are shown in Figure 15. At each visit over the 56 weeks period, the subjects with stable weight on placebo had their mean systolic blood pressure around the baseline value, while the mean systolic blood pressure for the subjects with stable weight on NB were almost always higher than the baseline value. .

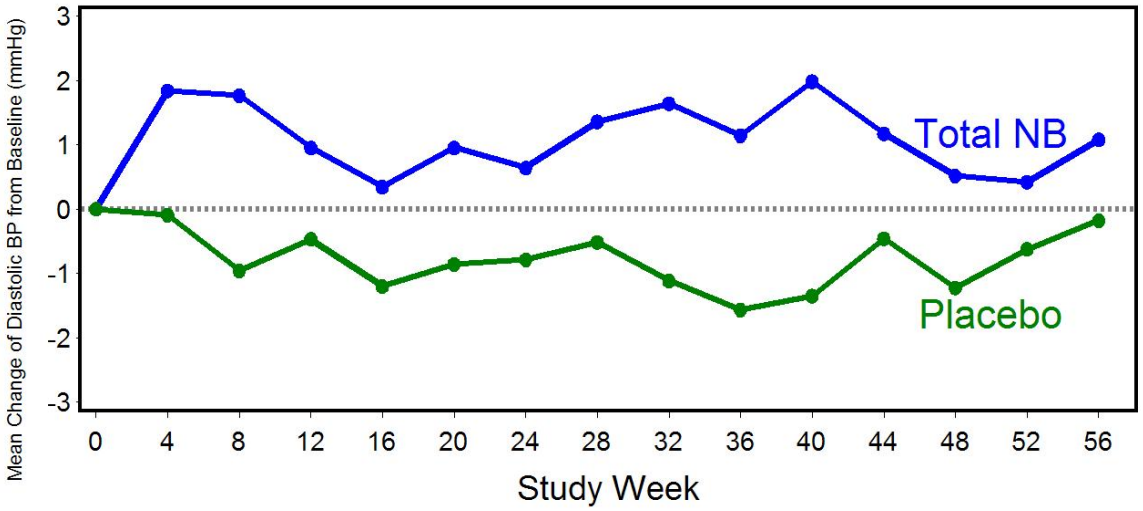
Figure 15: Mean Change of Systolic Blood Pressure by Treatment Group (Stable Weight Subpopulation)



Source: Created by reviewer.

Similarly, the mean changes of diastolic blood pressure from baseline to each visit over time are shown in Figure 16. Over the 56 weeks period, the subjects with stable weight on placebo had more beneficial change in diastolic blood pressure than the subjects with stable weight on NB. In the placebo group, the subjects with stable weight had their mean diastolic blood pressure below the baseline value. On the contrary, the subjects with stable weight in the NB groups had their mean diastolic blood pressure above the baseline value.

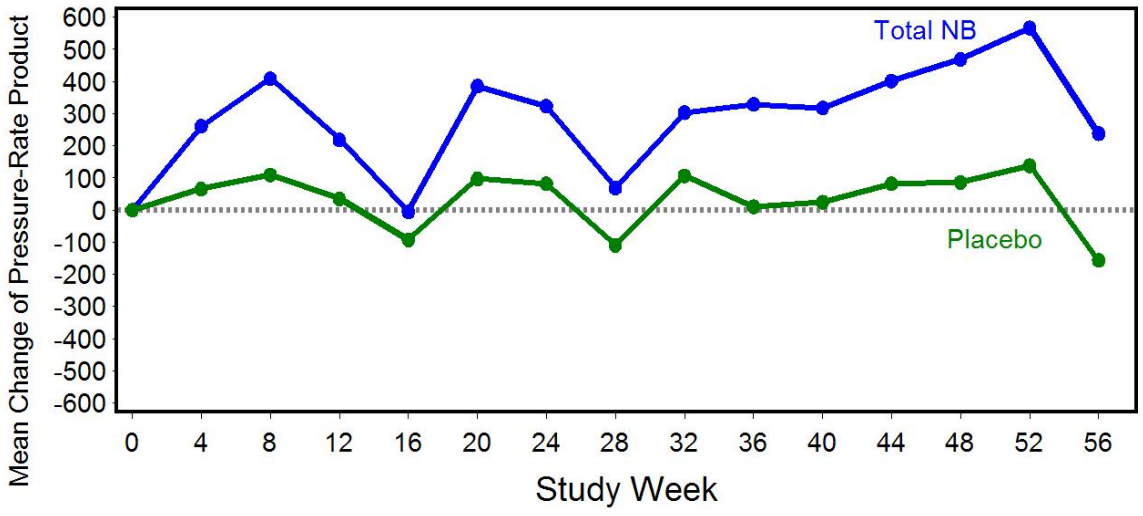
Figure 16: Mean Change of Diastolic Blood Pressure by Treatment Group (Stable Weight Subpopulation)



Source: Created by reviewer.

The mean changes of pressure-rate product from baseline to each visit over time are shown in Figure 17. Over the 56 weeks period, the subjects with stable weight on placebo had more beneficial change in pressure-rate product than the subjects with stable weight on NB.

Figure 17: Mean Change of Pressure-Rate Product by Treatment Group (Stable Weight Subpopulation)



Source: Created by reviewer.

5 Summary and Conclusions

5.1 Statistical Issues and Collective Evidence

As pre-specified in the applicant's Statistical Analysis Plan (SAP), phase 3 data from Studies NB-301, NB-302, NB-303, and NB-304 were pooled to evaluate the risk of treatment-emergent hypertension and the changes of vital signs. The comparison between the NB32 group and the placebo group is performed using the protocol-defined Cochran-Mantel-Haenszel (CMH) test with study as stratification factor for the binary outcome, and the protocol-defined analysis of covariance (ANCOVA) model for continuous outcomes.

Compared to placebo, the NB32 regimen was found to be associated with an increased risk of developing treatment-emergent hypertension. The incidence of treatment-emergent hypertension was higher in the NB32 group than in the placebo group (6.0% versus 4.1%). The pooled relative risk ratio between NB32 and placebo was 1.41 with a 95% confidence interval of (1.05, 1.91). In addition to occurring at a higher frequency, treatment-emergent hypertension also appeared to occur earlier in the NB32 group than in the placebo group. The onset time of treatment-emergent hypertension was statistically significantly earlier for NB32 as compared to placebo, with p-value=0.0039 for stratified log-rank test. The stratified Cox proportional hazard ratio (HR) was 1.57 with 95% CI (1.15, 2.13). Detailed analysis results are provided in Section 3.1.6.1.

As compared to the placebo group, the mean change of pulse rate from baseline to week 56 was statistically significantly higher in the NB32 group, in Study NB-301, NB-302, and NB-303 separately, as well as in the pooled analysis of all four studies. Among the four phase 3 studies, the difference between NB32 and placebo in mean change of pulse rate ranged from 0.89 to 1.21 bpm indicating a higher pulse rate for NB32 treated subjects. The pooled mean difference between NB32 and placebo was 1.11 bpm, with a 95% confidence interval of (0.62, 1.60). Detailed analysis results are provided in Section 3.1.6.2.1.

As compared to the placebo group, the mean change of systolic blood pressure from baseline to week 56 was statistically significantly higher in the NB32 group, in Study NB-301 and NB-302 separately, as well as in the pooled analysis of all four studies. The difference between NB32 and placebo in mean change of systolic blood pressure ranged from 0.67 to 2.63 mm Hg among the four studies indicating a higher systolic blood pressure in NB32 treated subjects. The pooled mean difference between NB32 and placebo was 1.41 mm Hg, with a 95% confidence interval of (0.76, 2.05). Detailed analysis results are provided in Section 3.1.6.2.1.

As compared to the placebo group, the mean change of diastolic blood pressure from baseline to week 56 was statistically significantly higher in the NB32 group, in Study NB-301 and NB-302 separately, as well as in the pooled analysis of all four studies. The difference between NB32 and placebo in mean change of diastolic blood pressure ranged from -0.01 to 1.32 mm Hg among the four studies indicating a higher diastolic blood

pressure in NB32 treated subjects. The pooled mean difference between NB32 and placebo was 0.58 mm Hg, with a 95% confidence interval of (0.13, 1.03). Detailed analysis results are provided in Section 3.1.6.2.1.

In addition to the change from baseline to week 56, the changes of vital signs over time from baseline to each visit during the 56 weeks of study follow-up were also assessed in this review. In the pooled analysis of all four phase 3 studies, the changes of vital sign were shown to be higher in the NB32 group than in the placebo group throughout the 56 weeks of study follow up. The difference between NB32 and placebo appeared to be largest at early time points, and tended to decrease at later study visits. At week 56, the pooled difference between NB32 and placebo was 1.25 bpm in pulse rate, 1.30 mm Hg in systolic blood pressure, and 0.38 mm Hg in diastolic blood pressure, while the pooled difference at week 8 was 2.35 bpm, 2.31 mm Hg, and 1.98 mm Hg respectively. Detailed analysis results are provided in Section 3.1.6.2.2.

Based on several exploratory analyses which incorporate whether a subject responded to treatment or not (responders=subjects who lost at least 5% of baseline body weight by the end of week 56), the placebo responders appeared to have the most beneficial change in all vital signs from baseline to week 56. On average, the subjects in the placebo responder group had a reduction of 2.09 bpm in pulse rate, 4.85 mm Hg in systolic blood pressure, and 3.82 mm Hg in diastolic blood pressure. On the contrary, NB32 non-responders had the least beneficial change of vital signs from baseline to week 56. On average, the subjects in the NB32 non-responder group had an increase of 0.51 bpm and 0.80 mm Hg in pulse rate and systolic blood pressure respectively and only had a reduction of 0.06 mm Hg in diastolic blood pressure. More details of this assessment can be found in Section 3.1.6.2.3.

5.2 Conclusions and Recommendations

Based on the pooled analysis of the four phase 3 randomized placebo-controlled clinical studies NB-301, NB-302, NB-303, and NB-304, the NB32 treatment regimen (Naltrexone 32 mg Sustained Release/Bupropion 360 mg Sustained Release) was found to be associated with an increased risk of developing treatment-emergent hypertension as compared to the placebo regimen. The incidence of treatment-emergent hypertension was statistically significantly higher at the nominal α=0.05 level in the NB32 group than in the placebo group (*p=0.02*). Furthermore, the onset time of treatment-emergent hypertension was also statistically significantly earlier in the NB32 group than in the placebo group (*p=0.004*).

Compared to placebo, the NB32 treatment regimen was demonstrated to have less of a beneficial effect on vital signs, measured as change of pulse rate, systolic and diastolic blood pressure, from baseline to week 56. At 56 weeks, the mean change of pulse rate from baseline was statistically significantly higher in the NB32 group than in the placebo group, in three out of the four phase 3 studies separately, as well as in the pooled analysis of all four studies (0.91 bpm versus -0.20 bpm, with *p<0.0001*). The mean change of systolic blood pressure from baseline to 56 weeks was statistically significantly higher in

the NB32 group than in the placebo group, in two out of the four phase 3 studies separately, as well as in the pooled analysis of all four studies (-0.22 mm Hg versus -1.62 mm Hg, with $p<0.0001$). Similarly, the mean change of diastolic blood pressure from baseline to 56 weeks was statistically significantly higher in the NB32 group than in the placebo group, in two out of the four phase 3 studies separately, as well as in the pooled analysis of all four studies (-0.76 mm Hg versus -1.35 mm Hg, with $p=0.1$).

Several exploratory analyses were done to explore the change of vital signs for treatment response groups (subjects who lost at least 5% of baseline body weight by the end of week 56 were identified as treatment responder). In the exploratory analyses, the placebo responders appeared to have the most beneficial change in vital signs (i.e. a reduction in the vital signs from baseline to week 56), while the NB32 non-responders tended to have the least beneficial change in vital signs.

Signatures/Distribution List

Primary Statistical Reviewer: Xiao Ding, PhD
Date: 12/21/2010

Concurring Reviewer(s):
Statistical Team Leader (Acting): Mat Soukup, PhD

Biometrics Division Director: Aloka Chakravarty, PhD

cc:

Archival NDA 200063
DMEP (HFD-510)/Meghna Jairath, Pharm.D
DMEP (HFD-510)/Pat Madara, MS
DMEP (HFD-510)/Eileen Craig, MD
DMEP (HFD-510)/Eric Colman, MD
DB7 (HFD-700)/Xiao Ding, PhD
DB7 (HFD-700)/Mat Soukup, PhD
DB7 (HFD-700)/Mark Levenson, PhD
DB7 (HFD-700)/Aloka Chakravarty, PhD
OB/IO (HFD-700)/Lillian Patrician, MS

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

/s/

XIAO DING
01/04/2011

MATTHEW J SOUKUP
01/04/2011
Concur with review

ALOKA G CHAKRAVARTY
01/04/2011

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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	200063/0
Drug Name:	Contrave™ (naltrexone / bupropion) extended-release tablets
Indication(s):	Weight management
Applicant:	Orexigen
Date(s):	Submission date: March 31, 2010
	PDUFA Goal Date: January 31, 2011
	Advisory Committee Date: December 7, 2010
Review Priority:	Standard
Biometrics Division:	Division of Biometrics 2
Statistical Reviewer:	Janice Derr, Ph.D.
Concurring Reviewers:	J. Todd Sahlroot, Team Leader and Deputy Division Director
	Thomas J. Permutt, Director, Division of Biometrics 2
Medical Division:	Division of Metabolism and Endocrinology Products
Clinical Team:	Eileen Craig, M.D., Medical Reviewer
	Eric Colman, M.D., Medical Team Leader and Deputy Division Director
	Mary H. Parks, M.D., Division Director
Project Manager:	Meghna Jairath

Keywords: clinical studies, NDA review

Table of Contents

1. EXECUTIVE SUMMARY 5

1.1 Conclusions and Recommendations 5

1.2 Brief Overview of Clinical Studies 7

1.3 Statistical Issues and Findings 8

2. INTRODUCTION 11

2.1 Overview 11

2.2 Scope of Statistical Review: Pivotal Efficacy and Safety Studies 12

2.2 Data Sources 19

3. STATISTICAL EVALUATION 20

3.1 Evaluation of Efficacy 20

3.2 Evaluation of Safety 62

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS 64

4.1 Sex, Race, Ethnicity and Age 64

4.2 Other Special/Subgroup Populations Defined by Medical Conditions at Baseline 65

4.3 Subgroups Defined by Responder Status 73

5. SUMMARY AND CONCLUSIONS 79

5.1 Statistical Issues and Collective Evidence 79

5.2 Conclusions 79

5.3 Recommendations for Labeling 80

LIST OF TABLES

TABLE 1 Phase 3 studies: Body weight (kg), percent change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF) 5
TABLE 2 Phase 3 studies; Body weight, proportion 5% responders at week 56 (week 28 for Study NB-303); Full Analysis Set (LOCF) 5
TABLE 3 Naltrexone / bupropion and placebo dosing daily during the titration and maintenance of the four Phase 3 studies 15
TABLE 4 Number of subjects, planned and actual, at randomization, in the FAS database and at completion (56 weeks) of the four Phase 3 studies 17
TABLE 5 Statistical power with respect to the continuous endpoint in the Phase 3 studies 19
TABLE 6 Data sources for studies 19
TABLE 7 Study NB-201 disposition for the primary treatment period (weeks 1 to 24), randomized population 20
TABLE 8 Study NB-201 primary and post-hoc evaluation of the Bup-Nal combinations at week 24 23
TABLE 9 Disposition in the four Phase 3 studies 26
TABLE 10 Subject demographic and baseline characteristics in the randomized subjects in each of the four Phase 3 studies 36
TABLE 11 Analysis populations defined for Studies NB-301, NB-302, NB-303 and NB-304 39
TABLE 12 Study NB-301; Body weight (kg), percent change from baseline to week 56 endpoint; Primary analysis and sensitivity analyses 44
TABLE 13 Study NB-301; Body weight, proportion 5% responders at week 56; primary analysis and sensitivity analyses. 45
TABLE 14 Study NB-302; Body weight (kg), percent change from baseline to week 56 endpoint; Primary analysis and sensitivity analyses 46
TABLE 15 Study NB-302; Body weight, proportion 5% responders at week 56; primary analysis and sensitivity analyses. 46
TABLE 16 Study NB-303; Body weight (kg), percent change from baseline to week 28 endpoint; Primary analysis and sensitivity analyses 47
TABLE 17 Study NB-303 at week 28; Body weight, proportion 5% responders at week 28; primary analysis and sensitivity analyses 47
TABLE 18 Study NB-304; Body weight (kg), percent change from baseline to week 56 endpoint; Primary analysis and sensitivity analyses 48
TABLE 19 Study NB-304; Body weight, proportion 5% responders at week 56; primary analysis and sensitivity analyses. 48
TABLE 20 Key secondary efficacy endpoints; results of analysis, showing pre-specified order for each Phase 3 study 56
TABLE 21 Phase 3 studies; Body weight, proportion of 10% responders at week 56 (week 28 for Study NB-303); Full Analysis Set (LOCF) 59
TABLE 22 Phase 3 studies: Waist circumference (cm), change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF) 59
TABLE 23 Phase 3 studies: Fasting HDL cholesterol (mg/dL), change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF) 60
TABLE 24 Phase 3 studies: Fasting triglycerides, % change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF) 60
TABLE 25 Phase 3 studies: Fasting triglycerides (mg/dL), change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF) 61
TABLE 26 Phase 3 studies: Impact of weight on quality of life (IWQOL-Lite) total score, change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF) 61

TABLE 27 Study NB-304 (diabetic subjects): HbA1c, change from baseline to week 56 endpoint Full Analysis Set (LOCF) 62
TABLE 28 Phase 3 studies: Systolic blood pressure (mm Hg), change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF) 63
TABLE 29 Phase 3 studies: Diastolic blood pressure (mm Hg), change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF) 63
TABLE 30 Proposed Patient Management Algorithm from Orexigen 77

LIST OF FIGURES

FIGURE 1 Disposition by week on study; Kaplan-Meier plots for Study NB-301, NB-302, NB-303 and NB-304 30
FIGURE 2 Study NB-301; For subjects who completed 56 weeks of study medication: mean weight change at each study visit. For subjects who discontinued: mean weight change at the study visit prior to discontinuation 32
FIGURE 3 Study NB-302; For subjects who completed 56 weeks of study medication, mean weight change at each study visit. For subjects who discontinued, mean weight change at the study visit prior to discontinuation 33
FIGURE 4 Study NB-303; For subjects who completed 56 weeks of study medication, mean weight change at each study visit. For subjects who discontinued, mean weight change at the study visit prior to discontinuation 34
FIGURE 5 Study NB-304; For subjects who completed 56 weeks of study medication, mean weight change at each study visit. For subjects who discontinued, mean weight change at the study visit prior to discontinuation 35
FIGURE 6 Study NB-301; Distribution of weight change at week 52; FAS analysis set 49
FIGURE 7 Study NB-302; Distribution of weight change at week 56; FAS analysis set 50
FIGURE 8 Study NB-303; Distribution of weight change at week 28; FAS analysis set 51
FIGURE 9 Study NB-304; Distribution of weight change at week 56; FAS analysis set 52
FIGURE 10 Blood pressure (mm Hg), repeated measures analysis of change from baseline to each visit: Primary safety dataset, double blind treatment phase 64
FIGURE 11 The proportion of 5% responders in subgroups (FAS/LOCF): Gender, race and ethnicity 66
FIGURE 12 The proportion of 5% responders in subgroups defined by medical condition at baseline (FAS/LOCF) 69
FIGURE 13 Systolic blood pressure (mm Hg), change from baseline to week 56 endpoint for all Phase 3 studies (pooled) by weight loss category 74
FIGURE 14 Systolic blood pressure (mm Hg), change from baseline to week 56 endpoint for all Phase 3 studies (pooled) by 5% responder category, completers population 75
FIGURE 15 Weight (kg), change from baseline to week 56 endpoint for all Phase 3 studies (pooled) by 5% responder category, completers population 75
FIGURE 16 Proposed patient management algorithm 78

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Confirmation of efficacy: The results of four Phase 3 studies are consistent and confirm the efficacy of naltrexone 32 mg /bupropion 360 mg (NB32) compared to placebo after 56 weeks of treatment in three studies and 28 weeks of treatment in one study. The co-primary endpoints were average weight loss compared to baseline and the percentage of subjects who lost at least 5% of baseline body weight. Results of alternate analysis models and other versions of the analysis population were consistent with the results from the primary analysis.

TABLE 1 Phase 3 studies: Body weight (kg), percent change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF)

	Group	n	Baseline mean (SD)	Percent Change from Baseline LSMean (SE)	LS Mean Difference from Placebo (95% CI)	p
Study NB-301 (week 56)	Placebo	511	99.3 (14.3)	-1.3 (0.3)		
Customary diet and	NB16	471	100.1 (14.4)	-5.9 (0.3)	-3.7 (-4.5, -2.9)	<0.001
behavior counseling	NB32	471	100.2 (16.3)	-6.1 (0.3)	-4.8 (-5.6, -4.0)	<0.001
Study NB-302 (week 56)	Placebo	193	101.9 (15.0)	-5.1 (0.6)		
Intensive lifestyle modification counseling	NB32	482	100.7 (15.4)	-9.3 (0.4)	-4.2 (-5.6, -2.9)	<0.001
Study NB-303 (week 28)	Placebo	456	99.3 (16.0)	-1.9 (0.3)		
Customary diet and behavior counseling	NB32	825	100.7 (16.7)	-6.5 (0.2)	-4.6 (-5.2, -3.9)	<0.001
Study NB-304 (week 56)	Placebo	159	105.0 (17.1)	-1.8 (0.4)		
Obese subjects with type 2 diabetes	NB32	265	106.5 (19.1)	-5.0 (0.3)	-3.3 (-4.3, -2.2)	<0.001

TABLE 2 Phase 3 studies; Body weight, proportion 5% responders at week 56 (week 28 for Study NB-303); Full Analysis Set (LOCF)

	Group	n	5% responders n(%)	Odds Ratio (vs. placebo)	95% CI of Odds Ratio	p
Study NB-301 (week 56)	Placebo	511	84 (16.4%)			
Customary diet and	NB16	471	186 (39.5%)	3.4	(2.5, 4.6)	<0.001
behavior counseling	NB32	471	226 (48.0%)	4.9	(3.6, 6.6)	<0.001
Study NB-302 (week 56)	Placebo	193	82 (42.5%)			
Intensive lifestyle modification counseling	NB32	482	320 (66.4%)	2.8	(2.0, 4.1)	<0.001
Study NB-303 (week 28)	Placebo	456	80 (17.5%)			
Customary diet and behavior counseling	NB32	825	459 (55.6%)	6.6	(5.0, 8.8)	<0.001
Study NB-304 (week 56)	Placebo	159	30 (18.9%)			
Obese subjects with type 2 diabetes	NB32	265	118 (44.5%)	3.4	(2.2, 5.5)	<0.001

Other key findings that support the efficacy of Contrave:

1. The results from the analysis of key secondary efficacy endpoints, such as triglycerides, HDL-cholesterol, waist circumference, and the total score on the Impact of Weight on the Quality of Life questionnaire, support the efficacy of the NB32 compared to placebo.

2. The placebo-adjusted effect of NB32 on the proportion of 5% responders was fairly similar among subgroups defined by BMI at baseline, or by the presence or absence of hypertension, dyslipidemia, or metabolic syndrome at baseline in each study.

3. Although the majority of subjects were female, white/Caucasian and not Hispanic, the studies were large enough to assess the effect of gender, white/Caucasian compared to black/African American, and Hispanic compared to not Hispanic on the efficacy of NB32. The placebo-adjusted effect of NB32 was fairly similar across these demographic factors in each of the Phase 3 studies.

4. In the study that enrolled subjects with Type 2 diabetes, the mean HbA1c at week 52 was in the direction of improved glycemic control in the NB32 group compared to the placebo group.

Considerations that may limit the efficacy of Contrave:

1. Early withdrawals from study medication: A substantial percentage of randomized subjects in each study and study arm, between 41% and 51%, discontinued study drug prior to week 56. Nausea associated with the NB combination contributed to the tendency to discontinue study medication, as did insufficient weight loss. For this reason, I believe that the 5% responder endpoint, with early withdrawals classified as non-responders, may reasonably capture the performance of NB32 in the target population of overweight and obese subjects who might use this drug. This adjustment brings in more non-responders and reduces the overall percentage of 5% responders with NB32 compared to placebo.

2. Increased mean systolic and diastolic blood pressure in NB32 compared to placebo: Based on the action of bupropion, it appears that to some extent, NB32 counteracts the beneficial relationship between weight loss and an improvement in blood pressure. These exploratory findings raise a concern about the overall balance of risk and benefit of Contrave.

Recommendations:

1. Proposed label: Recommendations for the proposed label are included in part 5.3 of this review.

2. Proposed patient management algorithm: The applicant has proposed an algorithm for the early identification of subjects who may not benefit from Contrave because of insufficient weight loss and/or an unacceptable increase in blood pressure. Recommendations for developing the statistical aspects of this algorithm and for submitting this information for further review are included in part 4.3.2 of this review.

3. Weight Management Guidance (2007 draft): The process of reviewing this application has identified several statistical issues that could usefully be addressed in the ongoing revision of the weight management guidance. These recommendations are summarized below:

    a. For a combination product, I recommend a statistical review of pivotal studies of the contribution of the components to the combination. I also recommend that at least one of the Phase 3 studies include single component arms as well as the combination and placebo arms.

    b. The large percentage of subjects who drop out is an acknowledged issue in weight loss studies. Nevertheless, I recommend that the guidance should emphasize the critical importance of tracking all randomized subjects in the Phase 3 studies, even those who discontinue from study medication and those who stop attending clinic visits.

    c. In a study with a large percentage of dropouts, the yes/no classification of weight-loss responders which classifies all discontinuations/dropouts as non-responders may best represent the effect of the drug in the intended target population.

    d. I recommend that the weight loss benchmarks for clinical significance be re-written for greater clarity in interpreting study results with respect to the benchmarks. For the continuous endpoint, I recommend the following revision: "The difference in mean weight loss between the active-product and placebo-treated groups is statistically significantly greater than 5%." For the benchmark for the categorical endpoint, I recommend a similar statement that is based on the odds ratio. I note that an odds ratio of 2.5 or greater meets the heuristic criteria for the categorical endpoint that are currently in the guidance.

    e. I recommend that both co-primary endpoints should be statistically significant at the 0.05 level in order to support the efficacy of the product. I believe that this revision would align the weight management guidance with the ICH-E9 guidance. The ICH-E9 guidance advises that in the event that a protocol identifies more than one primary endpoint, "the effect on the Type I error should be explained because of the potential for multiplicity problems …; the method of controlling Type I error should be given in the protocol."[1]

1.2 Brief Overview of Clinical Studies

The applicant used the results from the Phase 2 studies, particularly Study NB-201, to support the contribution of naltrexone and bupropion separately to the overall efficacy of the combination product. This was acceptable to the Division and follows the advice in the *Guidance for Industry: Developing Products for Weight Management* (2007 draft). A statistical review of Study NB-201 is included in part 3.1.1 of this review.

[1] Part II.B.5, *Guidance for Industry, E9 Statistical Principles for Clinical Trials,* September 1998

This NDA application includes the results from four Phase 3 studies. These studies evaluated the efficacy, safety and tolerability of naltrexone 32 mg /bupropion 360 mg and naltrexone 16 mg /bupropion 360 mg in obese and overweight subjects receiving customary diet and behavioral counseling, including prescribed exercise (Studies NB-301 and NB-303), in obese/overweight subjects undergoing intensive lifestyle modification counseling (Study NB-302), and in obese/overweight subjects with type 2 diabetes (Study NB-304). Studies NB-301, NB-302 and NB-303 enrolled subjects with a BMI $\geq$ 30 and $\leq$ 45 kg/m^2 for subjects with uncomplicated obesity, and with a BMI of $\geq$ 27 and $\leq$ 45 kg/m^2 for subjects with obesity and controlled hypertension and/or dyslipidemia. Study NB-304 enrolled obese/overweight subjects with type 2 diabetes (HbA1c > 7% and < 10%; not on injectable diabetes medications or inhaled insulin).

The number of subjects randomized in each study were as follows:

- Study NB-301: 1742 subjects were randomized; 578 to receive naltrexone 16 mg / bupropion 360 mg (NB16), 583 to receive naltrexone 32 mg / bupropion 360 mg (NB32) and 581 to receive placebo.

- Study NB-302: 793 subjects were randomized; 591 to receive NB32 and 202 to receive placebo.

- Study NB-303: 1496 subjects were randomized; 1001 to receive NB32 and 495 to receive placebo.

- Study NB-304: 505 subjects were randomized; 335 to receive NB32 and 170 to receive placebo.

Studies consisted of a screening period (up to 4 weeks), a 4-week titration period, and a 52-week maintenance period. Study visits occurred every 4 weeks. The primary endpoint for Study NB-303 occurred at week 28 rather than at week 56 because this study had a re-randomization protocol that started at week 28 and continued through week 44. During this period, subjects who were randomized to receive NB32 and who failed to achieve or maintain a 5% reduction in body weight were re-randomized to either continue treatment with NB32 or have their dose increased to naltrexone 48 mg /bupropion 360 mg. Because of this differential treatment of randomized arms, the Division requested that the primary endpoint be determined at week 28, with results from week 56 providing supportive information.

1.3 Statistical Issues and Findings

The following issues were important to the statistical review of this application:

1. <u>The Phase 2 study to confirm the contribution of each component to efficacy in the combination product</u>: The applicant used the results from the Phase 2 Study NB-201 to support the contribution of naltrexone and bupropion separately to the overall efficacy of Contrave. Based on the study results, the applicant compared the combination product to

placebo in the Phase 3 studies. This was acceptable to the Division and follows the advice in the *Guidance for Industry: Developing Products for Weight Management* (2007 draft).

However, in my opinion, the contribution of the bupropion and naltrexone components should have been confirmed in at least one of the Phase 3 studies. The reason for this opinion is that the pre-specified sequence of testing did not confirm the contribution of each component to the NB32 combination product in the Phase 2 study. The applicant used a second, post-hoc Bonferroni correction to conclude that each component contributed to the efficacy of the combination. While the p-value to which the Bonferroni correction was applied was fairly low and did not cause great concern, this experience supported the need for a more careful and timely statistical review of the pivotal study(ies) of the combination product prior to launching the Phase 3 studies. An additional evaluation of the combination in Phase 3 would have improved the strength of evidence supporting this pivotal evaluation. This could have been achieved by adding a bupropion arm and a naltrexone arm to one of the Phase 3 studies. I believe this aim could have been accomplished with a relatively small number of subjects in each component arm.

This review experience leads to my recommendation that a revision of the weight loss guidance should recommend (1) a review of the statistical analysis plan of the study(ies) used to confirm the contribution of each component of a combination drug; (2) a statistical review of the results of the pivotal combination study(ies); and (3) if a pivotal combination study is conducted at Phase 2, at least one Phase 3 study should include arms with the components as well as the combination arm and the placebo arm in order to confirm the combination in Phase 3.

2. Early discontinuation: A substantial percentage of randomized subjects in each study and study arm, between 41% and 51%, discontinued from taking study medication prior to week 56 (TABLE 9). Because many of these subjects discontinued before week 4, the first clinic visit, they were not included in the full analysis set (FAS), which was used for the primary efficacy analysis. However, I believe that these very early discontinuations provided information about Contrave that should not be lost. This is because the most frequently cited reason for early discontinuation from the naltrexone / bupropion combination was nausea, and the median time to discontinuation due to adverse events was 4 weeks. I believe it is reasonable to assume that subjects who dropped out this early were not likely to reach a 5% weight loss goal. In addition, subjects who discontinued early had lost less weight at the time of withdrawal, compared to the average weight loss at the same study week in subjects who completed the study (FIGURE 2-FIGURE 5). This trend was apparent in both the placebo and the naltrexone /bupropion arms in each study. For these reasons, I believe it is reasonable to classify subjects who discontinued early as non-responders on the 5% categorical endpoint. This permits the use of the intention-to-treat data base. In my opinion, this approach is most likely to represent the experience of the intended target population of Contrave. I note that the applicant did include this approach in the set of sensitivity analyses of each study, and that the placebo-adjusted effect of NB32 was statistically significant (TABLE 13, TABLE 15, TABLE 17 and TABLE 19). This adjustment brings in more non-responders and reduces the overall percentage of 5% responders with NB32 compared to placebo.

3. Benchmarks for clinical significance from the weight management guidance: Under the topic "Efficacy benchmarks," the weight management recommends, for the continuous endpoint, that "the difference in mean weight loss between the active-product and placebo-treated groups is at least 5% and the difference is statistically significant." In the four Phase 3 studies, the placebo-adjusted mean effect of NB32 was somewhat less than 5%, the benchmark for clinical significance, but the 95% confidence intervals include 5% (TABLE 12, TABLE 14, TABLE 16, TABLE 18). This raises a challenge to interpreting the guidance on this benchmark. In my opinion, an interpretation of the observed mean is not meaningful, because the observed mean is measured with error. In fact, the results from these studies demonstrate this issue. We can not conclude that the placebo-adjusted effect of NB32 was less than 5%. This is why I recommend re-wording this benchmark to refer to a statistically significant difference greater than 5%.

   For the categorical endpoint, I believe that the benchmark is unnecessarily complicated. The guidance states: "The proportion of subjects who lose greater than or equal to 5 percent of baseline body weight in the active-product group is at least 35 percent, is approximately double the proportion in the placebo-treated group, and the difference between groups is statistically significant." I believe that the intention behind these criteria is to make sure that a reasonable percentage of subjects are 5% responders with the active product, compared to placebo. However, I believe it is more useful to express the desired benchmarks in the target population in terms of the odds ratio. For example, when the percentage of active product responders is 35% and is twice that of the placebo group, the odds ratio is 2.5. When the percentage of active product responders is 50% and is twice that of the placebo group, the odds ratio is 3.0. This may define a reasonable working range for weight management products, from which to develop benchmarks. The odds ratio and confidence interval can be calculated from a logistic regression model that includes the main features of the study design, such as the treatment group assignment, stratification factors, and baseline covariates.

   I believe that the criteria for clinical significance would benefit from further consideration and development in the weight management guidance. In my opinion, expressing the benchmarks as statistical criteria provides a clearer means of interpreting them with respect to study results.

4. Controlling Type I error in key secondary efficacy endpoints: Within each study, secondary endpoints were tested in a pre-specified sequence, using a gate-keeping approach to control for Type I error. However, the long list of 15 to 20 key secondary endpoints, the differences among studies in the testing order, and the inclusion of additional endpoints in some studies but not others, has reduced the clarity of interpreting the significance of endpoints when viewed across studies. I believe that an improved approach to the control of Type I error in a list of secondary endpoints in a future application would impose more structure and consistency among studies, as follows: (1) identify groups of endpoints that are clinically related, such as the serum cholesterol endpoints; (2) within each clinical group, specify a testing order that is consistent from study to study; (3) use a method that allows for this structure and protects the overall Type I error among and within the clinical groups. With this more structured approach, it would be possible to add or remove clinical groups of

endpoints from study to study, while still maintaining clarity in interpreting results among studies.

5. The balance of risk and benefit in blood pressure and weight loss: Because one of the goals of weight loss is to reduce the risk of cardiovascular mortality, for which a reduction in blood pressure is an important biomarker, the Division requested additional exploratory evaluations of the relationship between weight loss and change in blood pressure in the placebo and Contrave groups. The summary findings are consistent with the interpretation that, to some extent, Contrave counteracts the beneficial relationship between weight loss and an improvement in blood pressure (FIGURE 13 - FIGURE 15). However, these findings need to be interpreted carefully, because weight loss is the outcome of the Phase 3 studies, and blood pressure is related both to weight loss and to bupropion.

   The applicant proposed a patient management algorithm for identifying patients who may not benefit from Contrave, either due to blood pressure elevation or insufficient weight loss or both (FIGURE 16). In this review, I provide recommendations for the additional analysis of Phase 3 study data in support of this algorithm (see part 4.3.2 of this review). My focus is on the development of prediction equations separately from each study, using the intention-to-treat population, with non-responder imputation for subjects who discontinued prior to the week at which the predictive value of early weight loss or early blood pressure changes are to be evaluated. Prediction equations can then be compared across studies and combined if they are reasonably similar.

2. INTRODUCTION

2.1 Overview

Contrave® (naltrexone HCL and bupropion HCL) Sustained-Release tablets is intended for the treatment of obesity, including weight loss and weight management. Contrave is a combination product of naltrexone and bupropion. Naltrexone is approved for the treatment of opiate and alcohol dependence. The usual adult dose is 50 mg/day for these indications. Bupropion is approved for the treatment of major depression and nicotine dependence. The usual adult dose is 300 mg/day for these indications. Bupropion is an aminoketone antidepressant with mixed dopamine and norepinephrine reuptake inhibitor action. Bupropion is known to cause weight loss when used for its currently approved indications. Naltrexone is a mu-opioid receptor antagonist. The applicant, Orexigen, based the development of the combination product on the hypothesis that the use of the two drugs in combination would lead to greater weight loss than would be seen with either naltrexone or bupropion alone[2].

The proposed daily dose of Contrave is 32 mg naltrexone / 360 mg bupropion. The applicant has developed two dosage strength tablets, Contrave 8/90 (naltrexone 8 mg / bupropion 90 mg) and Contrave 4/90 (naltrexone 4 mg / bupropion 90 mg). The tablets consist of a trilayer core that is composed of two drug layers containing the drug and excipients, and a more rapidly dissolving inert layer separating each drug. The applicant proposes an initial period of dose escalation,

[2] The source of this paragraph (paraphrased) is Part 2.2 (Introduction) of this NDA submission

starting with [redacted] (b) (4), if well tolerated, switching to Contrave 8/90 tablets over a 3-week period. The recommended daily dosage consists of two Contrave 8/90 tablets taken twice daily. [redacted] (b) (4)

Because both active ingredients are approved in the US for other indications, the Division has agreed that this is a 505(b)(2) NDA, with pertinent information from approved US prescribing information incorporated by reference.

The applicant proposes the following indication: "Contrave is indicated for the treatment of obesity and weight management, including weight loss and maintenance of weight loss, and should be used in conjunction with lifestyle modification. Contrave is recommended for patients with an initial body mass index $\geq$ 30 kg/m^2 or $\geq$ 27 kg/m^2 with one or more risk factors (e.g., diabetes, dyslipidemia, or hypertension."

2.2 Scope of Statistical Review: Pivotal Efficacy and Safety Studies

In this submission, the applicant describes the results from two Phase 2 and four Phase 3 studies to support the efficacy and safety of Contrave. The goals of the Phase 2 studies were to evaluate the efficacy, safety and tolerability of three fixed-dose combinations of naltrexone and bupropion; to evaluate the separate contribution of naltrexone and bupropion to the efficacy of the combination product; and to help determine which doses to evaluate in Phase 3. The Phase 3 studies all evaluated the combination product in comparison to placebo.

Because of the pivotal importance of Study NB-201 in evaluating the combination product, I am including a discussion of this study in this review.

2.2.1 Phase 2 studies:

The applicant used the results from the Phase 2 studies, particularly Study NB-201, to support the contribution of naltrexone and bupropion separately to the overall efficacy of Contrave. This is acceptable to the Division and follows the advice in the *Guidance for Industry: Developing Products for Weight Management* (2007 draft). The 2007 weight management guidance references the Code of Federal Regulations (CFR) concerning combination products, which states that two or more products may be combined into a single fixed-dosed combination when each component makes a contribution to the claimed effect or effects (21 CFR 300.50). The guidance recommends further that "the efficacy and safety of fixed dose combinations be compared with the individual product components of the combination and placebo in phase 2 trials of sufficient duration to compare the maximal or near-maximal weight-management effects of the placebo." The 2007 weight management guidance also states that "Once a fixed-dose combination has been deemed more effective than its individual components, the combination can then be examined versus placebo in phase 3 trials. This approach may preclude the need to include treatment groups for the individual components of the fixed-dose combination product in late-stage preapproval trials." This is the approach that the applicant took, with the concurrence of the Division.

The biometrics team reviewed the statistical analysis plan for Study NB-201 because of its pivotal contribution to the evaluation of the combination product (see review dated 5/31/06). The applicant submitted the results of Study NB-201 prior to initiating the Phase 3 studies, and the Division concurred that the results supported the contribution of each component towards the overall efficacy of the combination (see letter dated 12/7/06). I am including a brief review of the results of Study NB-201 in this review, both because of the pivotal nature of this study and because a statistical review of this study was not conducted at the time the results were submitted.

Study NB-201 was conducted in male and female patients, 18 to 60 years of age, with BMI ≥ 30 and ≤ 40 kg/m^2. Study drugs were titrated up during the first 4 weeks of blinded treatment to achieve the target dose for each treatment group. However, dose modifications and extended titration periods were permitted in the presence of adverse events and at the investigator's discretion. The primary endpoint was the weight at week 24 expressed as a difference from baseline and as a percentage change from the baseline weight. The study was conducted at 7 sites in the U.S, from 8/1/05 (first patient enrolled) to 12/13/06 (last patient completed). Dose groups, along with the daily dose of each component, are shown below:

- Bupropion 400 mg + Naltrexone 48 mg; n=67
- Bupropion 400 mg + Naltrexone 32 mg; n=70
- Bupropion 400 mg + Naltrexone 16 mg; n=67
- Bupropion 400 mg (monotherapy); n=66
- Naltrexone 48 mg (monotherapy); n=61
- Placebo; n= 88

Study NB-201 was conducted in two cohorts, with the combination arm with naltrexone 32 mg and additional patients to the placebo arm added in a second cohort after enrollment in the other five arms of the study was underway. The addition of the second cohort allowed for an evaluation of an intermediate dose of naltrexone. The statistical analysis plan was amended prior to unblinding of either cohort, to allow for the analysis of the added combination arm.

Study OT-101 was a proof of concept study with the primary objective of evaluating change in body weight between baseline and 16 weeks in two potential combination products and their associated monotherapy components (shown below):

- Bupropion 300 mg + Naltrexone 50 mg; n=60
- Naltrexone 50 mg; n=60
- Bupropion 300 mg; n=59
- Fluoxetine 60 mg + Naltrexone 50 mg; n=61
- Fluoxetine 60 mg; n=59
- Placebo; n=59

Study OT-101 was conducted in male and female patients, 18 to 60 years of age, with BMI ≥ 30 and ≤ 40 kg/m^2. The study consisted of a 4-week screening period, a 16-week primary treatment period, and an optional 32-week extension period. The duration of treatment after week 16 was contingent upon the results of the primary efficacy analysis at week 16. If this analysis indicated superior weight loss with one or more of the experimental treatment combinations, then the study was to proceed through week 48. However, if neither of the combination treatments had positive results, then the study was to be stopped after all subjects had completed week 24. Study OT-101 was conducted at 8 sites in the U.S., from 2/9/04 (first patient enrolled) until 11/1/04 (last patient completed).

2.2.2 Phase 3 studies

The statistical review covers the four Phase 3 studies that are described in this submission. These studies evaluated the efficacy, safety and tolerability of NB in obese and overweight subjects receiving customary diet and behavioral counseling, including prescribed exercise (Studies NB-301 and NB-303), in obese/overweight subjects undergoing intensive lifestyle modification counseling (Study NB-301), and in obese/overweight subjects with type 2 diabetes (Study NB-304). Studies NB-301, NB-302 and NB-303 enrolled subjects with a BMI ≥ 30 and ≤ 45 kg/m^2 for subjects with uncomplicated obesity, and with a BMI of ≥ 27 and ≤ 45 kg/m^2 for subjects with obesity and controlled hypertension and/or dyslipidemia. A more detailed description of each study is shown below:

- Study NB-301 was conducted in male and female subjects between 18 to 65 years of age with either uncomplicated obesity or with obesity/overweight with controlled hypertension and/or dyslipidemia. A total of 1742 subjects were randomized to receive naltrexone 16 mg / bupropion 360 mg (NB16; n=578), naltrexone 32 mg / bupropion 360 mg (NB32; n=583) or placebo (n=581). The study consisted of four periods: a screening period of up to 4 weeks (at least 2 visits), a titration period of 4 weeks (1 visit; TABLE 3); a study drug maintenance period of 52 weeks (14 visits), and a drug discontinuation period of 2 weeks (1 visit) for a total of 58 weeks of study duration. Subjects were to be seen every 4 weeks from baseline to week 56, and at week 58 following the 2-week drug discontinuation period. All subjects received ancillary therapy at baseline and at weeks 12, 24, 36 and 48. Ancillary therapy consisted of diet instruction and advice on behavior modification and exercise. Subjects who terminated study drug treatment before week 56 were encouraged to return for their scheduled visits to be weighed and for a waist circumference measurement at week 28 and week 56, as appropriate. The study was conducted at 34 sites in the U.S, from 10/4/07 (first patient enrolled) to 5/26/09 (last patient completed).

- Study NB-302 was conducted in male and female subjects between 18 to 65 years of age with either uncomplicated obesity or with obesity/overweight with controlled hypertension and/or dyslipidemia. A total of 793 subjects were randomized to receive NB32 (n=591) or placebo (n=202). The study consisted of a screening period (up to 4 weeks; TABLE 3), a 4-week titration period, and a 52-week maintenance period. Subjects also participated in an intense behavior modification program that included dietary instructions, twenty-eight 90-minute group sessions, and prescribed exercise. Subjects who prematurely discontinued study drug were encouraged to continue participation in the behavior modification program and to return to the study center for a body weight measurement (every 4 weeks) and waist circumference measurement (weeks 28 and 56). The study was conducted at 9 sites in the U.S, from 3/7/07 (first patient enrolled) to 9/12/08 (last patient completed).

- Study NB-303 was conducted in male and female subjects between 18 to 65 years of age with uncomplicated obesity or with obesity/overweight with controlled hypertension and/or dyslipidemia. There were 495 subjects randomized to receive placebo and 1001 randomized to receive NB32. The study consisted of a screening period (up to 4 weeks), a 4-week titration period, and a 52-week maintenance period. Subjects assigned to NB32 were randomized within each study center (1:1 ratio) to two alternative dosage schedules for

naltrexone, "fast" or "slow" (TABLE 3). For subjects in the fast titration group, the initial daily dose of naltrexone was 8 mg, which was increased by 8 mg each week until reaching 32 mg at week 4. For subjects in the slow titration group, the initial daily dose of naltrexone was 4 mg, which was increased by 4 mg each week until reaching 16 mg at week 4. The dose was increased to 32 mg at week 5. All subjects received ancillary therapy at baseline, weeks 12, 24, 36 and 48. Ancillary therapy consisted of diet instruction and advice on behavior modification and exercise.

Study visits occurred every 4 weeks. Beginning at week 28 through week 44, subjects randomized to receive NB32 who failed to achieve or maintain a 5% reduction in body weight were re-randomized to either continue treatment with NB32 or have their dose increased to naltrexone 48 mg / bupropion 360 mg (NB48). Subjects not re-randomized at Week 28 but who did not maintain at least 5% of baseline body weight loss during Weeks 32-44 were also re-randomized. Subjects were only re-randomized once. Subjects also received ancillary therapy consisting of dietary instruction and advice on behavior modification and exercise. The study was conducted at 36 sites in the U.S, from 12/6/07 (first patient enrolled) to 6/8/09 (last patient completed).

- Study NB-304 was conducted in obese/overweight subjects between 18 and 70 years of age with type 2 diabetes (HbA1c > 7% and < 10%; not on injectable diabetes medications or inhaled insulin). The safety and efficacy of a total daily dose of NB32 was compared to placebo. A total of 505 subjects were randomized to receive NB32 (n=335) or placebo (n=170). The study consisted of a screening period (up to 4 weeks), a 4-week titration period (TABLE 3), and a 52-week maintenance period. All subjects received ancillary therapy at baseline and weeks 4, 16, 28 and 40 consisting of diet instruction, advice on behavior modification, and physical activity suggestions. Subjects who prematurely discontinued study drug were encouraged to return to the study center for a body weight measurement (every 4 weeks) and waist circumference measurement (weeks 28 and 56). The study was conducted at 53 sites in the U.S, from 5/29/07 (first patient enrolled) to 6/1/09 (last patient completed).

TABLE 3 Naltrexone / bupropion and placebo dosing daily during the titration and maintenance of the four Phase 3 studies

Group	Studies	Titration Naltrexone (mg) / Bupropion (mg) Week 1 Days 1-7	Week 2 Days 8-14	Week 3 Days 15-21	Week 4 Days 22-28	Maintenance Weeks 5- 56
NB16	NB-301, NB-303	4 / 90	8 / 180	12 / 270	16 / 360	16 / 360
NB32fast	NB-301, NB-302, NB-303, NB-304	8 / 90	16 / 180	24 / 270	32 / 360	32 / 360
NB32slow	NB-303	4 / 90	8 / 180	12 / 270	16 / 360	32 / 360
Placebo	NB-301, NB-302, NB-303, NB-304	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0

Sources: Table 1 in each Clinical Study Reports, Study NB-301, Study NB-302, Study NB-303 and Study NB-304

Number of subjects in each trial: The applicant developed the size of each study to address: (1) the evaluation of efficacy from two co-primary endpoints; and (2) a general evaluation of

safety. A key resource was the *Guidance for Industry: Developing Products for Weight Management* (February 2007 draft). As part of my review, I also evaluated the size of each study with respect to the criteria for clinical significance, as described in the weight management guidance. I used a statistical interpretation of these criteria, and I note that this perspective is my own and does not reflect explicit statements in this guidance.

(1) For the evaluation of efficacy: Following the weight management guidance, a patient's body weight after one year of treatment in relation to baseline body weight is expressed in two different ways, as co-primary endpoints:

Continuous endpoint:	*the average weight loss at one year, expressed as a percentage change from baseline*
Categorical endpoint:	*the percentage of patients who lost at least 5% of their baseline body weight at one year*

The applicant used the following estimates and assumptions in calculating the number of subjects needed in each study for a statistical evaluation of efficacy:

- Approximately 40% of randomized subjects would discontinue from the study prior to the week 52 endpoint, including 20% who would not provide any post-baseline data (and therefore not be included in the intention-to-treat population).

- For the continuous endpoint:
  - A placebo-adjusted treatment effect of 5% (based on results from Study NB-201)
  - A standard deviation of 7% for studies NB-301 and NB-303 and 5% for studies NB-302 and NB-304. I note that these estimates are consistent with the results from study NB-201.

- For the categorical endpoint:
  - A placebo-adjusted treatment effect of 14% (50% of placebo-treated subjects and 64% of Contrave-treated subjects meeting the categorical endpoint) in studies NB-301, NB-301 and NB-303. The response rate for placebo was assumed to be similar to the response rates observed for the lifestyle modification alone arm, as reported by Wadden et al (2005)[3].
  - A placebo-adjusted treatment effect of 12.5% (15% of placebo subjects and 27.5% of Contrave-treated subjects in Study NB-304, based on a clinical study of rimonabant in overweight or obese patients with Type 2 diabetes (Scheen et al., 2006)[4]

- A two-tailed α of 0.05 and a target of at least 90% power

[3] Wadden TA, Berkowitz RI, Womble LG, Sarwer DB, Phelan S, Cato RK, et al. Randomized trial of lifestyle modification and pharmacotherapy for obesity. NEJM 2005; 353;20: 2111-20

[4] Scheen AJ, Finer N, Hollander P, Jensen MD, Van Gaal LF. Lancet 2006; 368: 1660-72.

(2) For the evaluation of safety: The weight management guidance suggests that "A reasonable estimation of the safety of a weight-management product upon which to base approval generally can be made when a total of approximately 3,000 subjects are randomized to active doses of the product and no fewer than 1,500 subjects are randomized to placebo for 1 year of treatment." The total number of subjects randomized in the four Phase 3 studies met these recommendations (TABLE 4). For the evaluation of safety, total exposure includes subjects in the Phase 2 studies as well.

TABLE 4 Number of subjects, planned and actual, at randomization, in the FAS database and at completion (56 weeks) of the four Phase 3 studies

Study	Treatment Arm	Number Randomized		Number in the FAS database		Number with at least 1-year exposure[2]	
		Planned	Actual	Predicted[1]	Actual	Predicted	Actual
NB-301	1. naltrexone 16mg / bupropion 360 mg	550	578	440	471	330	284
	2. naltrexone 32 mg / bupropion 360 mg	550	583	440	471	330	296
	3. placebo	550	581	440	511	330	290
NB-302	1. naltrexone 32 mg / bupropion 360 mg	600	591	480	482	360	342
	2. placebo	200	202	160	193	120	118
NB-303	1. naltrexone 32 mg / bupropion 360 mg	1000	1001	800	825	600	538
	2. placebo	500	495	400	456	300	267
NB-304	1. naltrexone 48 mg / bupropion 360 mg	350	335	280	265	210	175
	2. placebo	175	170	140	159	105	100
	Total for naltrexone / bupropion	3050	3088	2440	2514	1830	1635
	Total for placebo	1425	1448	1140	1319	855	775

Notes:
1 The predicted number in the FAS database is based on the assumption of 20% with no post-baseline data
2 The number predicted to have at least a 1-year exposure is based on the assumption of a 40% dropout rate
3 The actual number with at least a 1-year exposure is based on the number of subjects who completed the 56-week double-blind period while on study medication.
Sources: Summary by this reviewer, and TABLE 9, and TABLE 11 of this review

(3) Clinical significance of efficacy: Under the topic "Efficacy benchmarks," the guidance recommends:

> *In general, a product can be considered effective for weight management if after 1 year of treatment either of the following occurs:*
>
> - *The difference in mean weight loss between the active-product and placebo-treated groups is at least 5 percent and the difference is statistically significant*
>
> - *The proportion of subjects who lose greater than or equal to 5 percent of baseline body weight in the active-product group is at least 35 percent, is approximately double the proportion in the placebo-treated group, and the difference between groups is statistically significant.*

In my opinion, in order for these benchmarks to be interpretable, they need to be applied to the inference we can draw from the study results, not to the observed study results. For example, suppose the observed placebo-adjusted effect of Contrave were 5.1% for the continuous endpoint. Based on the observed difference, we would conclude in favor of Contrave. But suppose instead that the observed difference between means were 4.9%. Would we conclude against Contrave, or would this difference be close enough to 5% to decide in favor of Contrave? What would our conclusion be if the observed difference were 4.5%, or 4.0%, or 3.0%? Without understanding the variability of the difference observed from the study, these results are arbitrary and not interpretable. For this reason, I believe that it is much clearer to conclude in favor of Contrave if the 95% confidence interval of the difference between means excludes values less than 5%. This is the revised benchmark that I propose for the weight management guidance.

The Phase 3 studies had adequate statistical power to meet the benchmark of a statistically significant difference for the continuous endpoint (TABLE 5). The placebo-adjusted effect of Contrave was predicted to be 5% for the purpose of calculating statistical power. I note that in order to meet the revised benchmark for the continuous endpoint, the placebo-adjusted effect of the active product would have to be greater than 5%. This is because the revised benchmark of 5% is applied to the lower 95% confidence bound. For example, the Contrave studies would have had adequate statistical power to meet the revised benchmark, if the placebo-adjusted effect of Contrave had been predicted to be 6.5% or greater (TABLE 5).

With respect to the categorical endpoint, applying a statistical interpretation to the guidance benchmarks for clinical significance is more complicated. The description in the guidance actually covers three different ways to express and evaluate the categorical endpoint: as a single percentage (with a benchmark of 35%), as a comparison of the ratio of active product to placebo (with a benchmark of 2), and as a comparison of the difference between active product and placebo (with a benchmark of 0). I believe that the intention behind these criteria is to make sure that a reasonable percentage of subjects are 5% responders with the active product, compared to placebo. However, they seem unnecessarily complicated when comparing the percentage of 5% responders between active drug and placebo in the target population.

The benchmarks for the categorical endpoint do not appear to be very pertinent to the size of the Phase 3 studies for Contrave, for the following reasons:

- The percentage of responders to Contrave was predicted to be 64% for the nondiabetic subjects in studies NB-301, NB-301 and NB-303, which is well above the benchmark of 35%.
- The percentage of responders in the placebo was predicted to be 50% in these studies, making it unlikely that the ratio of approximately 2 would be met in these studies.
- The prediction for study NB-304 (diabetic subjects) was that the percentage of responders would be low and would not be likely to meet the benchmarks for clinical significance in the categorical endpoint.

TABLE 5 Statistical power with respect to the continuous endpoint in the Phase 3 studies

			Target population		
Study	Treatment Arm	Number Planned[1]	Predicted mean[2]	Statistical Power[3]	Using the revised benchmark, retrospective calculation of the minimum effect size[4]
NB-301	1. NB16	550	6%	99%	6.5%
	2. NB32	550	6%	99%	6.5%
	3. placebo	550	1%		---
NB-302	1. NB32	600	10%	99%	6.4%
	2. placebo	200	5%		---
NB-303	1. NB32	1000	6%	99%	6.3%
	2. placebo	500	1%		---
NB-304	1. NB32	350	6.5%	99%	6.5%
	2. placebo	175	1.5%		---

Notes:
[1] Expect 20% of randomized patients to have no post-baseline data
[2] Based on results from Study NB-201 and other considerations
[3] Based on a two-tailed $\alpha = 0.05$ and a standard deviation of 7%
[4] Applied retrospectively to the proposed revision to benchmark for clinical significance that difference between means should be $\geq 5\%$, and keeping other assumptions the same

2.2 Data Sources

The applicant submitted this NDA including the data to the FDA CDER Electronic Document Room (EDR). The submission is recorded in the EDR with the link shown in TABLE 6. Individual study reports were submitted for each study.

TABLE 6 Data sources for studies

Document: NDA 200063.0
CDER EDR link: \\CDSESUB1\N200063\
Company: Orexigen
Drug: Contrave
Submission date: March 31, 2010

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

The evaluation of efficacy starts with a review of the Phase 2 study NB-201, because this study evaluated the contribution of the naltrexone and bupropion components to the efficacy of the combination product. Following the review of study NB-201, I will review the evidence of efficacy from the four Phase 3 studies.

3.1.1 Phase 2 Combination Study

Disposition: The percentage of study dropouts ranged from 15.7% to 37.3% in the six arms of Study NB-201, with the highest percentage in the Bup+Nal48 combination arm (TABLE 7). The arms that included naltrexone were associated with a higher percentage of dropout than arms without naltrexone. The two arms with the highest percentage of dropouts (Bup+Nal48 and Bup+Nal16) also had the highest percentage of dropouts due to adverse events. The exception to this pattern is the Bup+Nal32 arm, which had the lowest percentage of dropouts in any arm (15.7). This is the arm that was added to the study as Cohort 2.

TABLE 7 Study NB-201 disposition for the primary treatment period (weeks 1 to 24), randomized population

	Treatment Group Number of subjects, n (%)					
	BPL + NPL N = 88	BPL + Nal 48 N = 61	Bup + NPL N = 66	Bup + Nal 16 N = 67	Bup + Nal 32 N = 70	Bup + Nal 48 N = 67
Study Disposition						
Completed	69 (78.4)	37 (60.7)	45 (68.2)	41 (61.2)	52 (74.3)	36 (53.7)
Terminated early[1]	16 (18.2)	19 (31.1)	15 (22.7)	23 (34.3)	11 (15.7)	25 (37.3)
Not treated	3 (3.4)	5 (8.2)	6 (9.1)	3[2] (4.5)	7 (10.0)	6 (9.0)
Primary reason for early termination						
Enrolled, but did not meet selection criteria	3 (3.4)	5 (8.2)	4 (6.1)	2 (3.0)	7 (10.0)	7 (10.4)
Failure to comply with protocol requirements	0 (0.0)	0 (0.0)	1 (1.5)	0 (0.0)	0 (0.0)	1 (1.5)
Subject moved	0 (0.0)	0 (0.0)	0 (0.0)	2 (3.0)	1 (1.4)	0 (0.0)
Insufficient weight loss	6 (6.8)	1 (1.6)	2 (3.0)	2 (3.0)	0 (0.0)	0 (0.0)
Adverse event	2 (2.3)	4 (6.6)	2 (3.0)	10 (14.9)	3 (4.3)	8 (11.9)
Withdrew consent	2 (2.3)	9 (14.8)	6 (9.1)	5 (7.5)	3 (4.3)	7 (10.4)
Lost to follow-up	6 (6.8)	3 (4.9)	4 (6.1)	5 (7.5)	4 (5.7)	7 (10.4)
Other	0 (0.0)	2 (3.3)	2 (3.0)	0 (0.0)	0 (0.0)	1 (1.5)
Time on study[3]						
0 weeks	0 (0.0)	1 (1.6)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
1 to 8 weeks	7 (8.0)	11 (18.0)	13 (19.7)	16 (23.9)	13 (18.6)	20 (29.9)
> 8 to 16 weeks	8 (9.1)	8 (13.1)	6 (9.1)	7 (10.4)	3 (4.3)	7 (10.4)
> 16 to 24 weeks	73 (83.0)	41 (67.2)	47 (71.2)	44 (65.7)	54 (77.1)	40 (59.7)

Source: Study NB-201 clinical report, Table 10-1

Analysis populations: The intention-to-treat (ITT) analysis set was defined as all subjects who were randomized and had at least one post-baseline body weight measurement. Missing values at study week 24 were imputed using the last non-missing observation carried forward (LOCF) method. An additional analysis population (Completers) consisted of subjects who remained in the study and were taking study medications at week 24.

Primary efficacy endpoint: The primary efficacy endpoint was the percent change from baseline in total body weight, measured between baseline and week 24.

Primary analysis model: The biometrics review team recommended a fully pre-specified analysis plan, because of the pivotal role of Study NB-201 in evaluating the combination therapy (see the letter to the sponsor dated June 14, 2006). Although the biometrics team recommended an analysis of covariance, with baseline body weight as a covariate and treatment group and center as factors in the model, the sponsor pre-specified an analysis of variance with treatment group as the main effect. The statistical analysis plan (SAP) also identified a closed testing sequence for the primary analysis, in the following order: (1) NB48 compared with placebo; (2) NB48 compared with naltrexone monotherapy; and (3) NB48 compared with bupropion monotherapy. Tests were evaluated in sequence, using a two-tailed α of 0.05. The comparisons of NB16 with placebo and monotherapy components were included as exploratory analyses, and the comparisons of NB32 with placebo and monotherapy components were not described in the SAP. This omission was likely due to the inclusion of the NB32 arm in the second cohort of the study. In my opinion, the pre-specified analysis plan was an acceptable evaluation of the NB48 combination.

Efficacy results: Following the pre-specified sequence of testing, we would conclude that the Bup-Nal 48 combination is not supported, because the contribution of the naltrexone 48 component is not statistically significant (TABLE 8, part 1, applicant's pre-specified method). The placebo-adjusted effect of the BN48 combination was an average loss of 3.5% of baseline weight at week 24. This is statistically significant from the placebo arm. The contribution of the bupropion monotherapy component, 3.1%, was statistically significant, but the contribution of the naltrexone 48 mg monotherapy component, 1.7%, was not statistically significant (TABLE 8).

The applicant speculated that "The relatively greater frequency of permanent dose modifications and early treatment discontinuations associated with the higher naltrexone dose used in the NB48 treatment group are likely offsetting potential weight loss effects in this analysis."[5] For this reason they conducted an additional, post-hoc analysis of all three of the Bup-Nal combinations. They evaluated each of the 9 pairwise comparisons involved in this post-hoc analysis at an adjusted α of $0.5/9 = 0.0056$. This post-hoc approach enabled them to evaluate the Bup-Nal 32 and the Bup-Nal 16 combinations. Using this approach, they concluded that the efficacy of both of these combinations was supported, because the monotherapy components were associated with a statistically significant contribution to the average weight loss of the combination (TABLE 8, part 2, post-hoc method).

On the basis of these results, the Division agreed with the applicant that the Phase 3 studies could be conducted with comparisons of NB combinations conducted only with placebo (see letter dated 12/7/06). This decision follows recommendations from the 2007 weight management guidance, which states that "Once a fixed-dose combination has been deemed more effective than its individual components, the combination can then be examined versus placebo in phase 3 trials. This approach may preclude the need to include treatment groups for the individual

[5] Study NB-201 report, part 11.4.1.1

components of the fixed-dose combination product in late-stage preapproval trials."[6] However, in my opinion, the contribution of the Bup and Nal 32 mg components should have been confirmed in at least one of the Phase 3 studies. The reasons for this opinion include: (1) The Bup-Nal 32 mg dose which was used in Phase 3 trials was added on as a second cohort in study NB-201 and was therefore not included in pre-specified analysis plan for evaluating the combination; (2) The Bup-Nal 32 mg dose combination was supported by a set of post-hoc hypothesis tests. Although the test-wise α was adjusted to allow for multiple hypothesis tests, the post-hoc nature of this evaluation raises questions about the repeatability of this finding. I believe that the Bup-Nal 32 mg dose combination could have been evaluated in a relatively small number of patients randomized to a bupropion monotherapy arm and a naltrexone 32 mg monotherapy arm within one of the Phase 3 studies. The evaluation could have taken place at 26 weeks rather than at the 52 week endpoint period. This approach would have provided a stronger confirmation of the combination product at the intended dosage.

[6] 2007 weight management guidance (draft), part IV D.

TABLE 8 Study NB-201 primary and post-hoc evaluation of the Bup-Nal combinations at week 24

Combination	vs.	Monotherapy or Placebo	N Comb	N Mono	Baseline mean weight (kg) ± SD Comb	Baseline mean weight (kg) ± SD Mono	Adjusted mean % change from baseline at Week 24 ± SE Comb	Adjusted mean % change from baseline at Week 24 ± SE Mono	Difference in adjusted mean change, Combination – Monotherapy or Placebo LS Mean (95% CI)	P-value
1. Applicant's pre-specified method[1]										
Step 1. BN48	vs.	Placebo	54	84	94.4 ± 13.0	97.9 ± 13.4	-4.3 ± 0.7	-0.8 ± 0.5	-3.5 (-5.2 to -1.9)	<0.0001
Step 2. BN48	vs.	N48		49		96.0 ± 10.6		-1.2 ± 0.7	-3.1 (-5.0 to -1.3)	0.0009
Step 3. BN48	vs.	B		57		98.5 ± 13.1		-2.7 ± 0.6	-1.7 (-3.4 to 0.1)	0.0684
	Decision: The BN48 combination is not supported because the contribution of the naltrexone 48 component is not statistically significant.									
2. Applicant's post-hoc method[2]										
BN48	vs.	Placebo	54	84	94.4 ± 13.0	97.9 ± 13.4	-4.3 ± 0.7	-0.8 ± 0.5	-3.5 (-5.2 to -1.9)	<0.0001
	vs.	N48		49		96.0 ± 10.6		-1.2 ± 0.7	-3.1 (-5.0 to -1.3)	0.0009
	vs.	B		57		98.5 ± 13.1		-2.7 ± 0.6	-1.7 (-3.4 to 0.1)	0.0684
	Decision: The BN48 combination is not supported because the contribution of the naltrexone 48 component is not statistically significant.									
BN32	vs.	Placebo	63	84	95.8 ± 12.6	97.9 ± 13.4	-5.4 ± 0.6	-0.8 ± 0.5	-4.7 (-6.2 to -3.1)	<0.0001
	vs.	N48		49		96.0 ± 10.6		-1.2 ± 0.7	-4.3 (-6.0 to -2.5)	<0.0001
	vs.	B		57		98.5 ± 13.1		-2.7 ± 0.6	-2.8 (-4.5 to -1.1)	0.0015
	Decision: The BN32 combination is supported because the contribution of both components is statistically significant.									
BN16	vs.	Placebo	54	84	95.3 ± 13.8	97.9 ± 13.4	-5.4 ± 0.7	-0.8 ± 0.5	-4.6 (-6.2 to -3.0)	<0.0001
	vs.	N48		49		96.0 ± 10.6		-1.2 ± 0.7	-4.2 (-6.1 to -2.4)	<0.0001
	vs.	B		57		98.5 ± 13.1		-2.7 ± 0.6	-2.7 (-4.5 to -1.0)	0.0026
	Decision: The BN16 combination is supported because the contribution of both components is statistically significant.									

Notes:
[1] Applicant's pre-specified method: Evaluate the three comparisons in the order shown, at a two-tailed α of 0.05. If all three comparisons are statistically significant, conclude that the BN48 combination is supported.
[2] Applicant's post-hoc method: Evaluate all 9 comparisons simultaneously at an adjusted two-tailed α of 0.05/9 = 0.0056.
Source: Study NB-201 clinical report, Table 5.1.1.1c

3.1.2 Phase 3 Studies

3.1.2.1 Subject disposition

Completing 56 weeks of treatment with study drug: The disposition event of interest in each Phase 3 study was the completion or withdrawal from 56 weeks of treatment with study drug. Study NB-303 also evaluated the completion of 28 weeks of study drug. Subjects were free to discontinue their participation in the study (i.e., withdraw consent) at any time and without prejudice to further treatment. The investigator was able to withdraw a subject because of a safety risk or adverse event. The study protocols listed reasons why a subject might have their study drug discontinued, including non-adherence to at least 70% of study drug for two consecutive months or discontinuation of study drug for any reason for a period of at least 15 consecutive days.

A substantial percentage of randomized subjects in each study and study arm, between 41% and 51%, discontinued from taking study medication prior to week 56 (TABLE 9). This percentage of subjects was fairly similar in the placebo arm and the combination arm(s) within each study (TABLE 9). A large percentage of early discontinuation is typical of weight loss studies. Investigators in this field have proposed and evaluated different ways to evaluate weight loss programs and/or drugs, given that a large percentage of subjects are likely to discontinue before the primary endpoint period.[7] The weight management guidance recommends estimating the effect of a drug by several different methods. This sensitivity analysis should reflect the time dynamics and reasons for early discontinuation.

The naltrexone/bupropion combination was associated with early discontinuation due to adverse events, more so than the placebo arm in each study (TABLE 9A). More subjects identified "adverse events" as their reason for discontinuing from the naltrexone/bupropion combination than any other reason. This finding is consistent across studies. Moreover, the median time to withdrawal due to adverse events was 4 weeks in each study, which is the end of the titration period (TABLE 9B). This was the earliest median time to withdrawal for any of the reasons for discontinuation. The slow titration schedule for the NB32 combination did not appear to delay the median time to withdrawal compared to the fast titration schedule (TABLE 3, TABLE 9B). Nausea was the most frequently cited reason for dropout due to adverse event in the naltrexone/bupropion arms, followed by headache and dizziness (TABLE 9C).

The re-randomization period in Study NB-303 which took place at weeks 28-44 involved 251 subjects in the NB32 group who had failed to achieve or maintain at least 5% body weight loss from baseline. Of these, 128 were assigned to continue with NB32, and 107 (83.5%) completed treatment; 123 were assigned to NB48 and 104 (84.6%) completed treatment. The most common reason for discontinuation of drug was loss to follow-up.

[7] For example, see Gadbury, GL, CS Coffee and DB Allison, 2003: Modern statistical methods for handling missing repeated measurements in obesity trial data: beyond LOCF. Obesity Reviews 4: 175-184.

On average, subjects who withdrew from study medication early had lost less weight at the time of withdrawal, compared to the average weight loss at the same study visit in subjects who completed 56 weeks of study medication (FIGURE 2-FIGURE 5). To assess this pattern, I calculated the mean weight change at each study visit in subjects who had discontinued study medication in the interval between a given study visit and the previous study visit. I compared this mean weight change to the weight change at the given study visit by patients who completed 56 weeks of study medication. The difference between subjects who completed and subjects who withdrew from study medication is apparent both in the combination arm(s) and the placebo arm in each study (FIGURE 2-FIGURE 5). The difference in means between completers and withdrawals at any given visit appears to be fairly constant from week 4 up through week 28, which is the time frame for most withdrawals. From week 28 to week 56, the difference in means at any given visit is more variable (FIGURE 2-FIGURE 5). This variability may reflect (1) the smaller numbers of patients who withdrew after week 28; (2) different reasons for withdrawing from study medication towards the end of a study rather than close to the start; (3) uncertainty in the definition of the status of subjects who withdrew from taking study medication close to the actual 56 week endpoint[8]. For this reason, the mean weight change from patients who discontinued from study medication is measured with the greatest precision from week 1 through 28, when most of the discontinuations took place.

Study visits following discontinuation from study drug: In the event that a subject discontinued study drug treatment prior to week 56, the protocol instructed the study investigator to make every effort to have the subject return as soon as possible for an early termination visit that included all assessments outlined for the week 56 visit. In addition, the investigator was instructed to encourage the subject to return for their scheduled visits to be weighed. The off-study drug weights were used in the intention-to-treat analysis population as a sensitivity analysis.

[8] See the notes in FIGURE 2-FIGURE 5, and additional discussion in the definition of the "completers" analysis population, in part 3.1.2.3 of this review

TABLE 9 Disposition in the four Phase 3 studies

Study NB-301 at week 56 Customary diet and behavioral counseling	**naltrexone 16mg / bupropion 360 mg**	**naltrexone 32 mg / bupropion 360 mg**	**placebo**	**Total**
A. Disposition				
Randomized	578	583	581	1742
Completed	284 (49.1%)	296 (50.8%)	290 (49.9%)	870 (49.9%)
Withdrawn:	294 (50.9%)	287 (49.2%)	291 (50.1%)	872 (50.1%)
Adverse events	*122 (21.1%)*	*112 (19.2%)*	*56 (9.6%)*	*290 (16.6%)*
Withdrew consent	*63 (10.9%)*	*60 (10.3%)*	*90 (15.5%)*	*213 (12.2%)*
Lost to follow-up	*76 (13.1%)*	*65 (11.1%)*	*66 (11.4%)*	*207 (11.9%)*
Insufficient weight loss	*12 (2.1%)*	*12 (2.1%)*	*40 (6.9%)*	*64 (3.7%)*
Other[1]	*21 (3.8%)*	*38 (6.5%)*	*39 (6.7%)*	*98 (5.6%)*
B. Median time on study (weeks) prior to withdrawal by reason for withdrawal				
Adverse events	*4*	*4*	*9*	
Withdrew consent	*13*	*12*	*12*	
Lost to follow-up	*7*	*5*	*9*	
Combined Other[2]	*12*	*23.5*	*14*	
C. Most frequently cited adverse events as the reason for withdrawal				
Nausea or Vomiting	*31*	*41*	*3*	
Headache	*9*	*5*	*4*	
Dizziness or Vertigo	*15*	*7*	*3*	

Table 9, continued **Study NB-302 at week 56** **Intensive program for behavior modification**	**naltrexone 32 mg / bupropion 360 mg**	**placebo**	**Total**
A. Disposition			
Randomized	591	202	793
Completed [3]	342 (57.9%)	118 (58.4%)	460 (58.0%)
Withdrawn:	249 (42.1%)	84 (41.6%)	333 (42.0%)
Adverse events	*150 (25.4%)*	*25 (12.4%)*	*175(22.1%)*
Withdrew consent	*43 (7.3%)*	*24 (11.9%)*	*67 (8.4%)*
Lost to follow-up	*22 (3.7%)*	*17 (8.4%)*	*39 (4.9%)*
Insufficient weight loss	*3 (0.5%)*	*6 (3.4%)*	*9 (1.1%)*
Other[1]	*31 (5.2%)*	*12 (5.9%)*	*43 (5.4%)*
B. Median time on study (weeks) prior to withdrawal by reason for withdrawal			
Adverse events	*4*	*15*	
Withdrew consent	*16*	*22.5*	
Lost to follow-up	*5*	*15*	
Combined Other[2]	*21*	*25*	
C. Most frequently cited adverse events as the reason for withdrawal			
Nausea or Vomiting	*31*	*0*	
Headache	*5*	*1*	
Dizziness or Vertigo	*5*	*0*	

Table 9, continued **Study NB-303 at weeks 28 and 56** **Customary diet and behavioral counseling**	**naltrexone 32 mg / bupropion 360 mg**	**placebo**	**Total**
A. Disposition			
Randomized	1001	495	1496
Completed at week 28	619 (61.8%)	319 (64.4%)	
Withdrawn by week 28	382 (38.2%)	176 (35.6%)	
Completed at week 56	538 (53.7%)	267 (53.9%)	805 (53.8%)
Withdrawn by week 56:	463 (46.3%)	228 (46.1%)	691 (46.2%)
Adverse events	*241 (24.1%)*	*68 (13.7%)*	*309 (20.7%)*
Withdrew consent	*75 (7.5%)*	*56 (11.3%)*	*131 (8.8%)*
Lost to follow-up	*77 (7.7%)*	*48 (9.7%)*	*125 (8.4%)*
Insufficient weight loss	*19 (1.9%)*	*33 (6.7%)*	*52 (3.5%)*
Other[1]	*51 (5.1%)*	*23 (4.6%)*	*74 (4.9%)*
B. Median time on study (weeks) prior to withdrawal by reason for withdrawal			
Titration[4]*:*	*Fast*	*Slow*	
Adverse events	*4*	*4*	*9*
Withdrew consent	*9*	*12.5*	*12.5*
Lost to follow-up	*8*	*12*	*9*
Combined Other[2]	*14*	*18.5*	*16*
C. Most frequently cited adverse events as the reason for withdrawal			
Nausea or Vomiting	*68*	*1*	
Headache	*26*	*4*	
Dizziness or Vertigo	*12*	*1*	

Table 9, continued **Study NB-304 at week 56** **Obese subjects with type 2 diabetes**	**naltrexone 32 mg / bupropion 360 mg**	**placebo**	**Total**
A. Disposition			
Randomized	335	170	505
Completed	175 (52.2%)	100 (58.8%)	275 (54.5%)
Withdrawn:	160 (47.8%)	70 (41.2%)	230 (45.5%)
Adverse events	*98 (29.3%)*	*26 (15.3%)*	*124 (24.6%)*
Withdrew consent	*21 (6.3%)*	*15 (8.8%)*	*36 (7.1%)*
Lost to follow-up	*22 (6.6%)*	*15 (8.8%)*	*37 (7.3%)*
Insufficient weight loss	*5(1.5%)*	*6 (3.5%)*	*11 (2.2%)*
Other[1]	*14 (4.2%)*	*8 (4.7%)*	*22 (4.4%)*
B. Median time on study (weeks) prior to withdrawal by reason for withdrawal			
Adverse events	*4*	*11*	
Withdrew consent	*17*	*30*	
Lost to follow-up	*5*	*13*	
Combined Other[2]	*15*	*17*	
C. Most frequently cited adverse events as the reason for withdrawal			
Nausea or Vomiting	*42*	*0*	
Headache	*6*	*0*	
Dizziness or Vertigo	*4*	*1*	

Notes:

[1] The "Other" category is a combination of categories identified by the applicant: drug non-compliance, failure to comply with protocol requirements, subject moved, subject became pregnant, enrolled but did not meet selection criteria, other primary reason, randomized but study drug not dispensed, death, and other primary reason not listed.

[2] The "Combined Other" category includes all of the categories from "Other" and "Insufficient weight loss."

[3] A discrepancy in the number per arm in the completers analysis set in Study NB-302 and the total number per arm classified as "completed" in the disposition analysis has been noted by this reviewer and by the applicant. This is attributed to the definition of "completer" that was used for the analysis population. For more information, see Part 3.1.2.3 of this review.

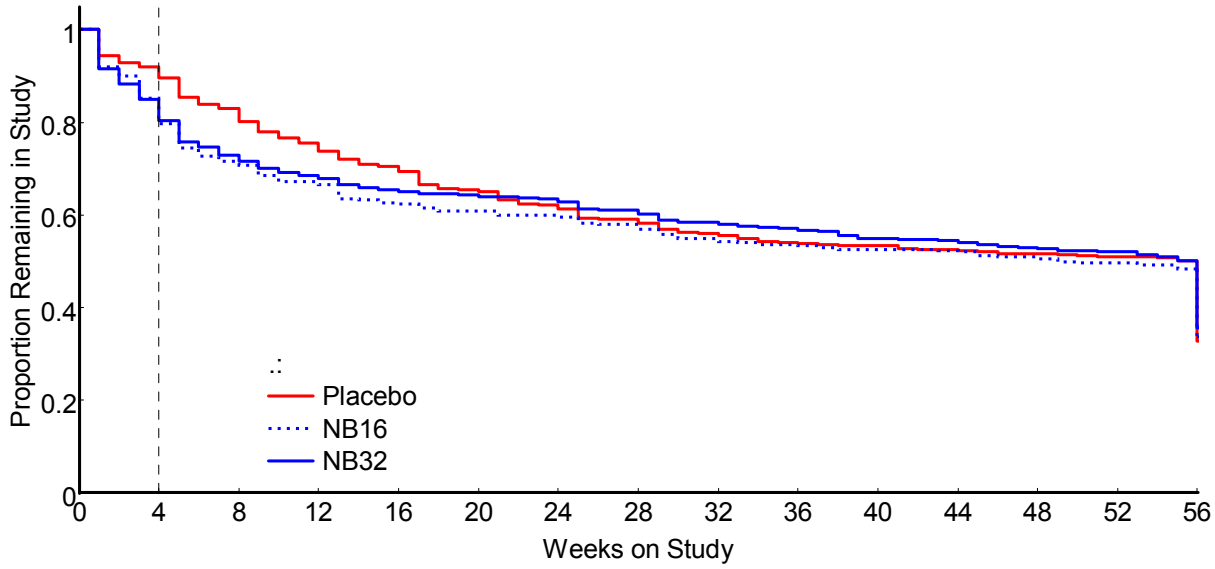
[4] Subjects assigned to NB32 in Study NB-303 were randomly allocated within center to two schedules of dose titration, "fast" and "slow" as described in Part 2.2.2 of this review.

Sources:

Tables ISE.301.101, ISE.302.1-1, ISE303.1-1 and ISE304.1-1, and analysis by this reviewer

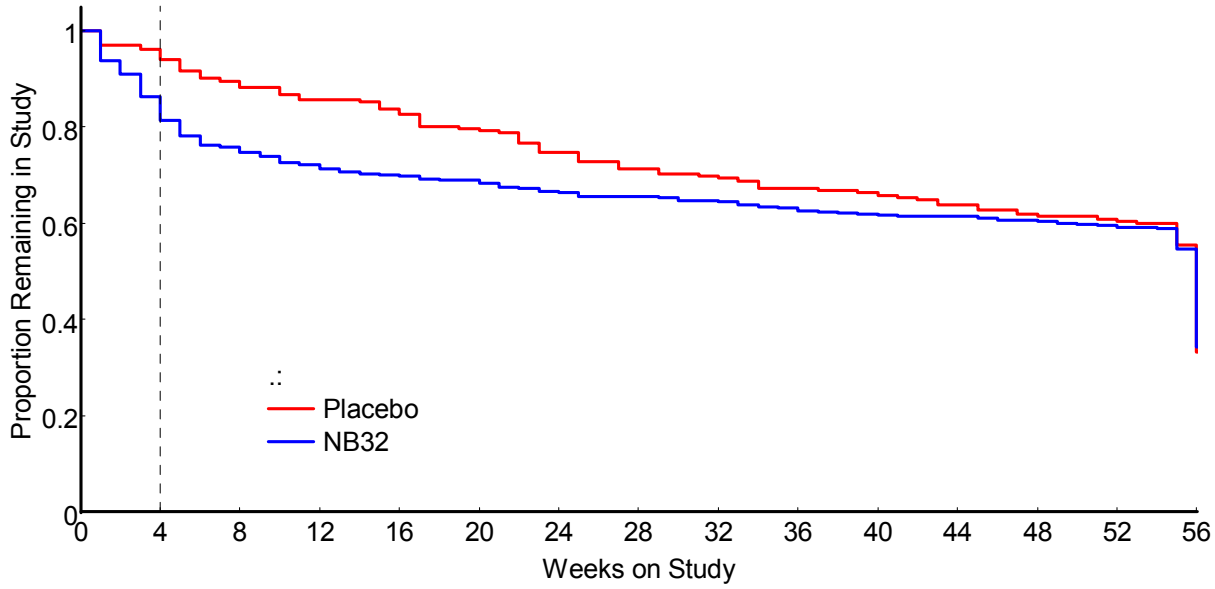
FIGURE 1 Disposition by week on study; Kaplan-Meier plots for Study NB-301, NB-302, NB-303 and NB-304

Study NB-301: Customary diet and behavioral counseling



Note for Study NB-301: The dashed gridline at week 4 represents the end of the drug titration period.

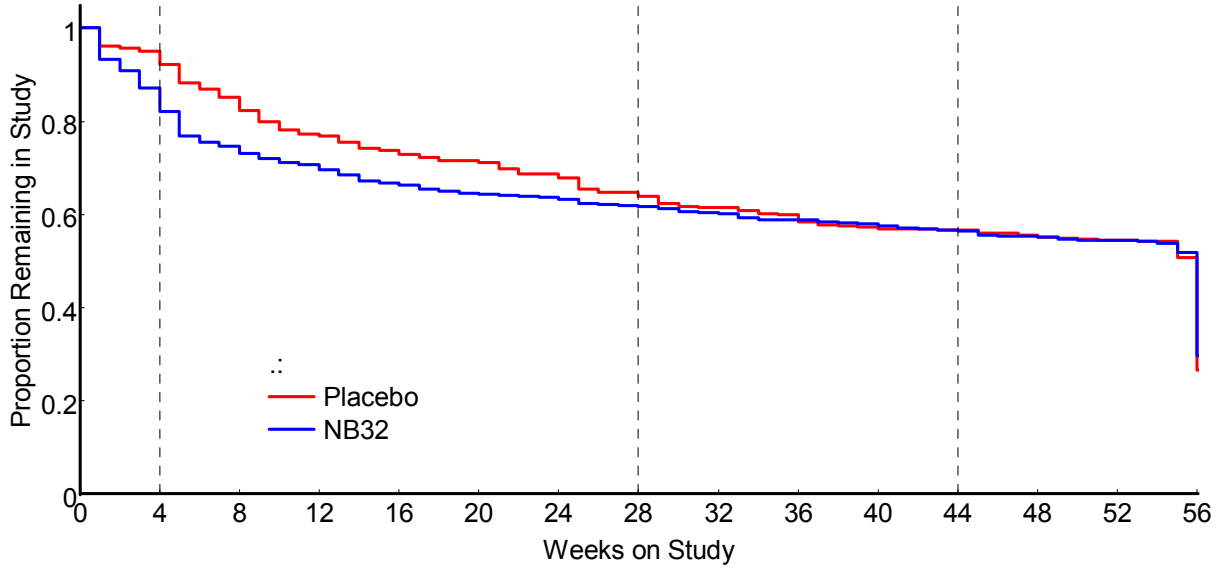
Study NB-302: Intensive program for behavior modification



Note for Study NB-302: The dashed gridline at week 4 represents the end of the drug titration period.

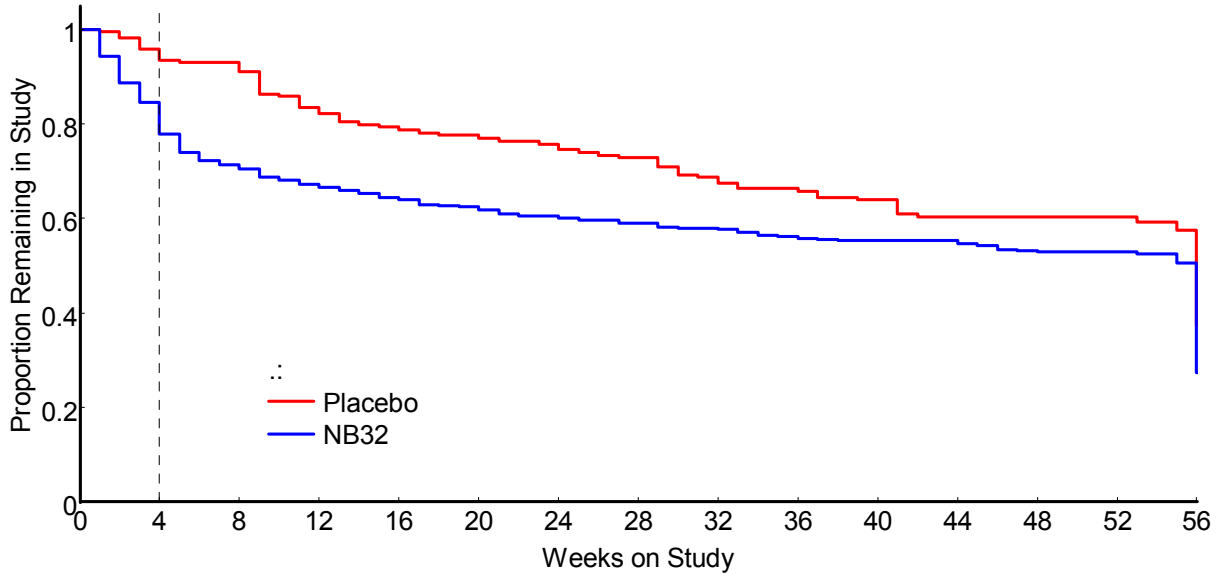
Figure 1, continued

Study NB-303: Customary diet and behavioral counseling



Notes for Study NB-303: The dashed gridline at week 4 represents the end of the drug titration period. From weeks 28-44, NB32-treated subjects who failed to achieve or maintain at least 5% of body weight loss from baseline were re-randomized (1:1 ratio) to continue NB32 or receive NB48. Subjects were re-randomized only once.

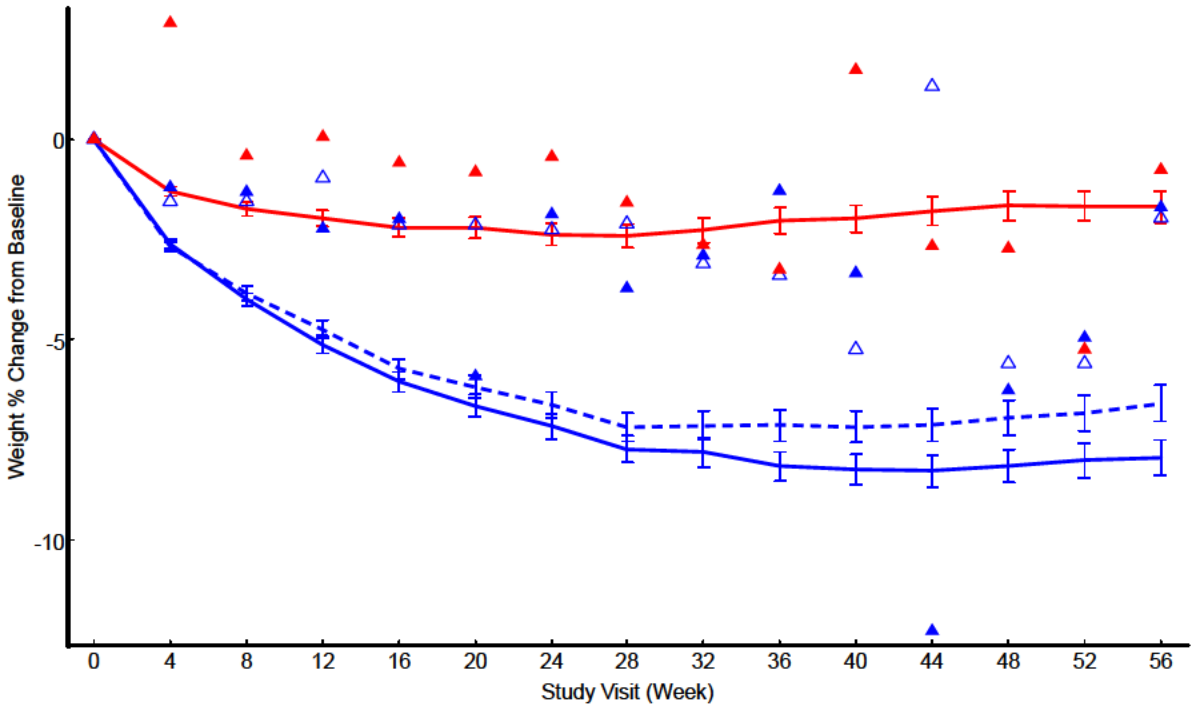
Study NB-304: Obese subjects with type 2 diabetes



Notes for Study NB-304: The dashed gridline at week 4 represents the end of the drug titration period.

Source: Analysis by this reviewer

FIGURE 2 Study NB-301; For subjects who completed 56 weeks of study medication: mean weight change at each study visit. For subjects who discontinued: mean weight change at the study visit prior to discontinuation

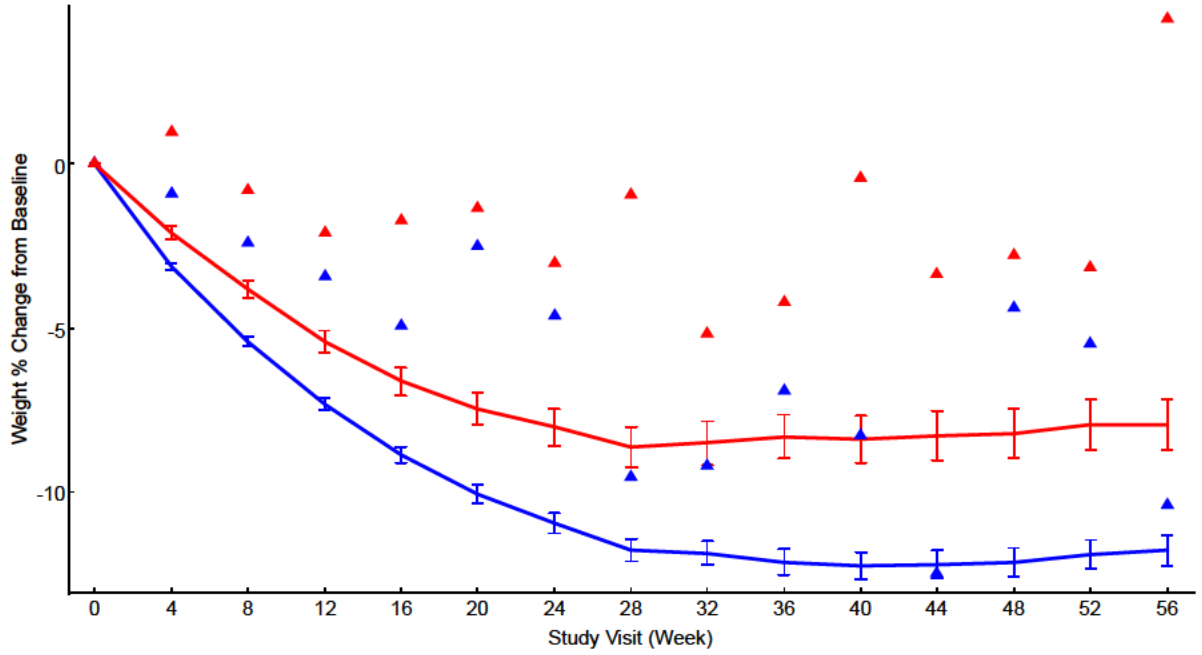


Notes:

- For subjects who stayed on study medication through week 56, the mean ± SEM is shown as a longitudinal profile. Subjects who completed study medication had treatment durations of 55-61 weeks in the efficacy database provided by the applicant (variable TRTDURWK).
- Single symbols represent the mean of subjects who discontinued in the time interval between the previous study visit (t-1) and the given study visit (t). A small number of subjects who were classified as "did not complete" study medications had treatment durations of 53-57 weeks: 3 in the NB16 arm, 8 in the NB32 arm, and 6 in the placebo group (variable TRTDURWK). This may represent uncertainty in the classification of subjects who withdrew from study medication close to the study endpoint.
- The dose titration period was week 0 through week 4.

Source: Analysis by this reviewer

FIGURE 3 Study NB-302; For subjects who completed 56 weeks of study medication, mean weight change at each study visit. For subjects who discontinued, mean weight change at the study visit prior to discontinuation

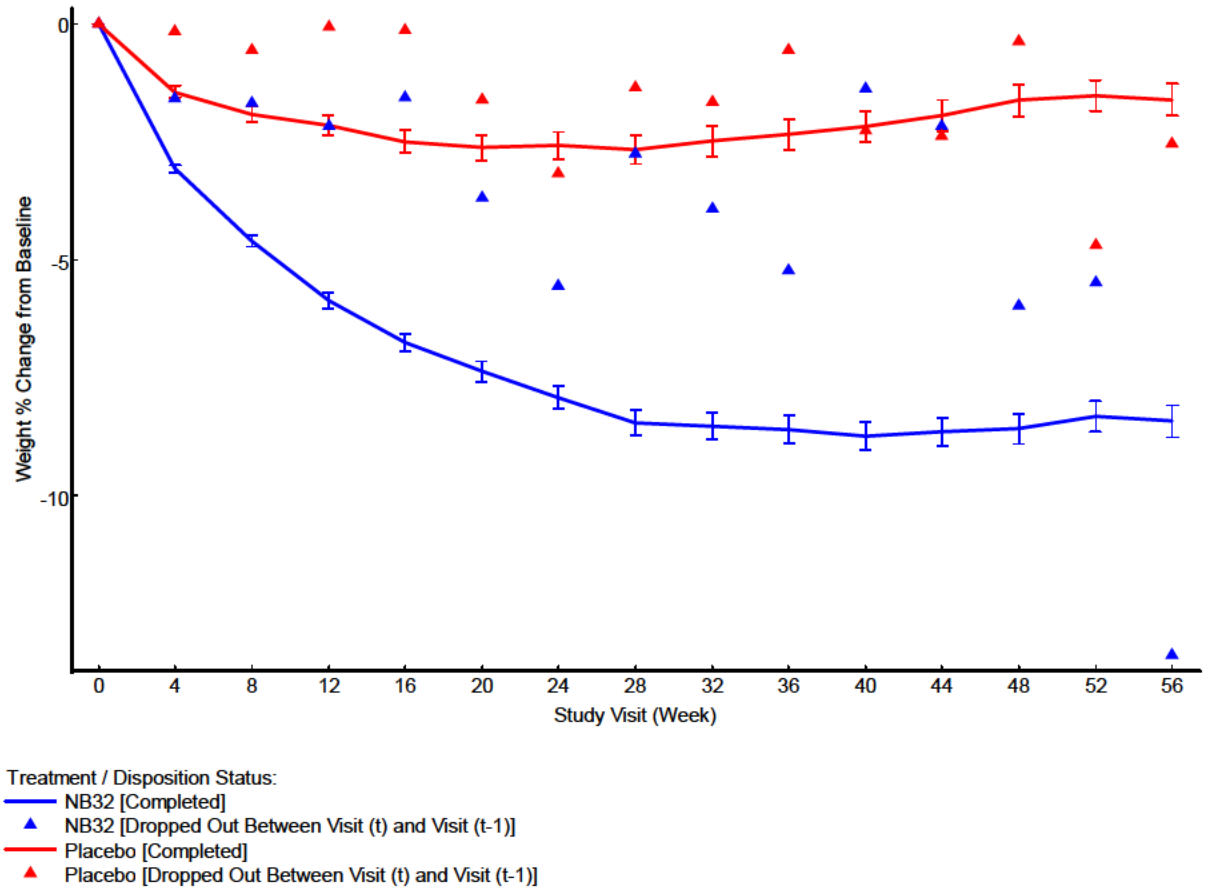


Notes:

- For subjects who stayed on study medication through week 56, the mean ± SEM is shown as a longitudinal profile. Subjects who completed study medication had treatment durations of 54-61 weeks in the efficacy database provided by the applicant (variable TRTDURWK).
- Single symbols represent the mean of subjects who discontinued in the time interval between the previous study visit (t-1) and the given study visit (t). Some subjects who were classified as "did not complete" study medications had treatment durations of 53-61 weeks: 7 in the NB32 arm, and 4 in the placebo group (variable TRTDURWK). This may represent uncertainty in the classification of subjects who withdrew from study medication close to the study endpoint.
- The dose titration period was week 0 through week 4.

Source: Analysis by this reviewer

FIGURE 4 Study NB-303; For subjects who completed 56 weeks of study medication, mean weight change at each study visit. For subjects who discontinued, mean weight change at the study visit prior to discontinuation

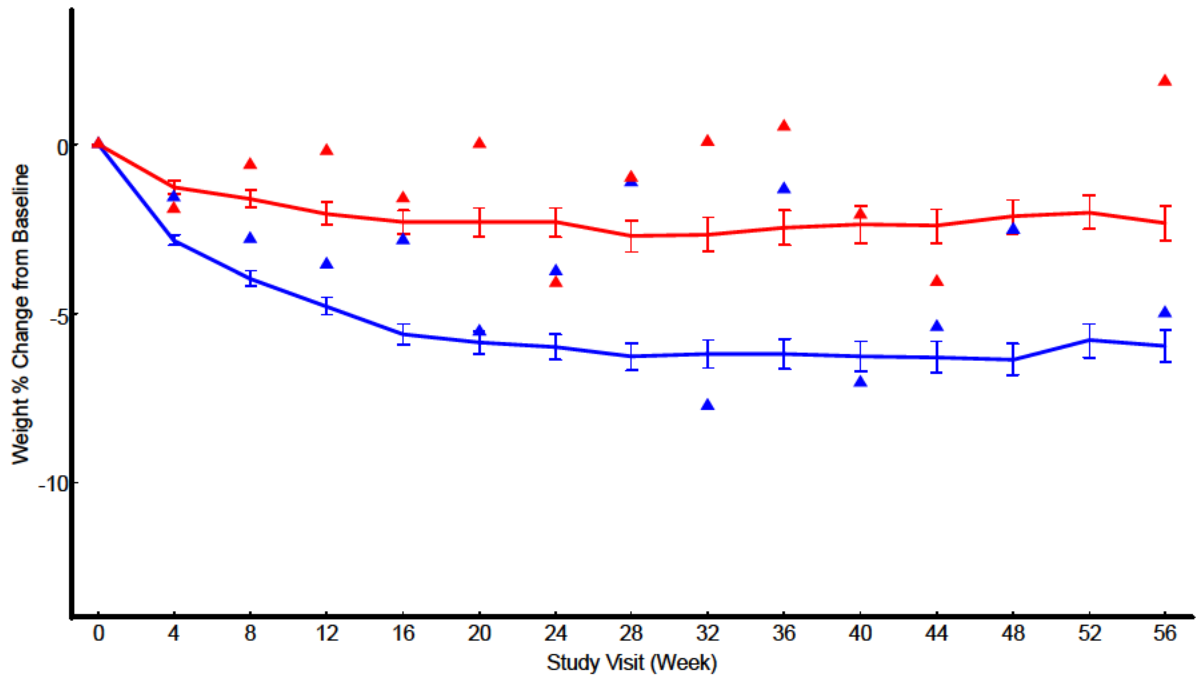


Notes:

- For subjects who stayed on study medication through week 56, the mean ± SEM is shown as a longitudinal profile. Subjects who completed study medication had treatment durations of 53-64 weeks in the efficacy database provided by the applicant (variable TRTDURWK).
- Single symbols represent the mean of subjects who discontinued in the time interval between the previous study visit (t-1) and the given study visit (t). A small number of subjects who were classified as "did not complete" study medications had treatment durations of 53-59 weeks: 7 in the NB32 arm, and 2 in the placebo group (variable TRTDURWK). This may represent uncertainty in the classification of subjects who withdrew from study medication close to the study endpoint.
- Means include subjects in the NB32 arm who were re-randomized to receive NB48 in the re-randomization period starting at week 28 and continuing through week 44 (with the re-randomized dose continued through week 56).
- The dose titration period was week 0 through week 4.

Source: Analysis by this reviewer

FIGURE 5 Study NB-304; For subjects who completed 56 weeks of study medication, mean weight change at each study visit. For subjects who discontinued, mean weight change at the study visit prior to discontinuation



Notes:

- For subjects who stayed on study medication through week 56, the mean ± SEM is shown as a longitudinal profile. Subjects who completed study medication had treatment durations of 55-62 weeks in the efficacy database provided by the applicant (variable TRTDURWK).
- Single symbols represent the mean of subjects who discontinued in the time interval between the previous study visit (t-1) and the given study visit (t). A small number of subjects who were classified as "did not complete" study medications had treatment durations of 53-55 weeks: 2 in the NB32 arm, and 2 in the placebo group (variable TRTDURWK). This may represent uncertainty in the classification of subjects who withdrew from study medication close to the study endpoint.
- The dose titration period was week 0 through week 4.

Source: Analysis by this reviewer

3.1.2.2 Subject demographic and baseline characteristics

Differences and similarities among the Phase 3 studies reflected the enrollment of subjects with type 2 diabetes in Study NB-304 and subjects who did not have type 2 diabetes in Studies NB-301, NB-302 and NB-303 (TABLE 10). The average age of diabetic subjects in Study NB-304 was greater than the average of subjects in the other three studies. Males and females were approximately equally represented among diabetic subjects in Study NB-304, whereas the large

majority of subjects in the other three studies were female. A greater percentage of diabetic subjects in Study NB-304 had hypertension, dyslipidemia, and/or metabolic syndrome at baseline than did the non-diabetic subjects in the other three studies. All four studies were relatively similar in the distribution of racial groups, with the large majority of subjects from the Caucasian racial group (TABLE 10). Average weight at baseline was also relatively similar at approximately 100 kg. A similar percentage, approximately 60%, of subjects had BMI $\geq$ 35 kg/m^2 at baseline in all four studies.

TABLE 10 Subject demographic and baseline characteristics in the randomized subjects in each of the four Phase 3 studies

	Study NB-301 n=1742	**Study NB-302 n=793**	**Study NB-303 n=1496**	**Study NB-304 n=505**
Age (years)				
Mean ± SD	44.1 ± 11.2	48.5 ± 10.7	44.3 ± 11.2	53.8 ± 9.3
Median	45.0	47	45.0	54.0
Range	18 to 66	19 to 65	18 to 65	20 to 72
≥ 65 years (n, %)	---	---	---	61 (12.1%)
Sex				
Male (n, %)	260 (14.9%)	80 (10.1%)	229 (15.3%)	220 (43.6)
Female (n, %)	1482 (85.1%)	713 (89.9%)	1267 (84.7%)	285 (56.4)
Race[1]				
Caucasian/ White	1307 (75.0%)	554 (69.9%)	1249 (83.5%)	401 (79.4%)
African American/ Black	338 (19.4%)	189 (23.8%)	205 (13.7%)	81 (16.0%)
Asian	14 (0.8%)	8 (1.0%)	16 (1.1%)	12 (2.4%)
Native Hawaiian / Pacific Islander	15 (0.9%)	1 (0.1%)	4 (0.3%)	1 (0.2%)
American Indian / Alaska Native	48 (2.8%)	8 (1.0%)	12 (0.8%)	4 (0.8%)
Other	20 (1.1%)	33 (4.2%)	10 (0.7%)	6 (1.2%)
Ethnicity[2]				
Hispanic/ Latino	229 (13.1%)	77 (9.7%)	114 (7.6%)	58 (11.5%)
Not Hispanic/ Latino	1513 (86.9%)	716 (90.3%)	1382 (92.4%)	447 (88.5%)
Weight (kg)				
Mean ± SD	99.6 ± 15.0	100.6 ± 15.3	100.0 ± 16.3	104.5 ± 18.3
Median	98	100.0	98.0	104.0
Range	62 to 155	66 to 162	66 to 168	64 to 167
BMI (kg/m^2)				
Mean ± SD	36.2 ± 4.2	36.5 ± 4.2	36.2 ± 4.4	36.4 ± 4.7
Median	36.0	36.0	36.0	36.0
Range	27.0 to 47.0	28.0 to 46.0	27.0 to 46.0	27.0 to 46.0
Obesity class (n, %)				
BMI < kg/m^2	39 (2.2%)	9 (1.1%)	39 (2.6%)	29 (5.7%)
BMI ≥ 30 and < 35	659 (37.8%)	271 (34.2%)	584 (39.0%)	160 (31.7%)
BMI ≥ 35 and < 40	633 (36.3%)	309 (39.0%)	507 (33.9%)	174 (34.5%)
BMI ≥ 40	411 (23.6%)	204 (25.7%)	366 (24.5%)	142 (28.1%)
Subgroups (n, %)				
Hypertension	315 (21.7%)	123 (15.5%)	318 (21.3%)	315 (62.4%)

	Study NB-301 n=1742	Study NB-302 n=793	Study NB-303 n=1496	Study NB-304 n=505
Dyslipidemia	736 (50.7%)	351 (44.3%)	823 (55.0%)	425 (84.2%)
Metabolic	388 (26.7%)	188 (23.7%)	450 (30.1%)	358 (70.9%)
Impaired fasting glucose	376 (25.9%)	177 (22.3%)	407 (27.2%)	---[1]
HbA1c > 8.0%	---[2]	---	---	181 (35.8%)
Sulfonylurea pharmacotherapy	---	---	---	241 (52.3%)

Notes:
[1] Impaired fasting glucose was not a subgroup for the diabetic population in Study NB-304
[2] HbA1c and Sulfonylurea pharmacotherapy did not define subgroups for the non-diabetic populations of Studies NB-301, NB-302 and NB-303.
Sources:
Tables ISE.301.1-3, ISE.301.1-5, ISE 302.1-3, ISE 302.1-5, ISE 303.1-3, ISE 303.1-5, ISE 304.1-3, ISE 304.1-5

3.1.2.3 Analysis populations

All four Phase 3 studies used the same definitions for the analysis populations, with exceptions as described below:

Full Analysis Set (FAS): The full analysis set included all subjects who were randomized, had a baseline body weight measurement, and at least one post-baseline body weight measurement while on study drug. Baseline was defined as the last non-missing measurement across all the visits before or at the time of randomization. Endpoint was defined as the last non-missing postbaseline measurement while on study drug (last observation carried forward, LOCF). Efficacy assessments that occurred within 1 day following the last dose date were considered valid.

Per Protocol Analysis Set (PP): The per protocol analysis set included all subjects in the FAS who received at least 28 weeks of study treatment, were "at least 70% compliant" with study medication.

Intent-to-Treat Analysis Set (ITT): The ITT analysis set included all randomized subjects with a baseline and postbaseline body weight, where endpoint was defined as the last non-missing postbaseline (LOCF_ITT) measurement during the double-blind treatment phase (irrespective of being on study drug at the time of the last measurement). A key difference between the FAS and the ITT set is that for the ITT analysis the efficacy assessments were considered valid even if they occurred while off study drug.

Completers Analysis Set: For studies NB-301, NB-303 and NB-304, the completers analysis set included all randomized subjects with a baseline measurement, a postbaseline body weight measurement, and who completed 56 weeks of treatment. For study NB-302, the definition of the completers analysis set was somewhat different from the other studies. This difference, and a discrepancy between the number of subjects classified as "completers" by disposition status and

the number in the completers database led to a request for clarification from the applicant. The applicant provided the following explanation[9]:

> Study NB-302 was the first completed Phase 3 study in the clinical development program. The definition of the Completers Analysis Set for that study had been pre-specified in the statistical analysis plan prior to database lock. Upon evaluation of the study results, Orexigen noted the same discrepancy identified by the Agency. Orexigen determined that the requirement for a subject to have their Week 56 study visit while on study medication in order to be qualified as a "Completer" appeared to be the source of this discrepancy. Specifically, it was noted that all of these subjects had completed the study but were not included in the "Completers Analysis Set" because they had discontinued study drug more than 1 day prior to the Week 56 study visit. Therefore, it was decided that a definition for completers that more closely matched the study treatment disposition results was deemed to be more appropriate. For this reason, the definition of completers was modified for Studies NB-301, NB-303 and NB-304 prior to study completion and database lock. The new definition dropped the requirement that the Week 56 measurement occur "while on study drug" and only required that the subject complete study participation. The completer classification was applied consistently to both placebo and NB subjects within each study.

For Study NB-303, an additional completers analysis set was defined for week 28.

Safety Analysis Set: The safety analysis set for analysis during the double blind treatment phase included all randomized subjects who were administered at least one tablet of study treatment and had at least one investigator context / assessment at any time after the start of study treatment, regardless of whether or not they discontinued the study.

[9] See the applicant's response to the Division's questions received by Orexigen on October 27, 2010

TABLE 11 Analysis populations defined for Studies NB-301, NB-302, NB-303 and NB-304

Analysis set, n (%)	Study NB-301			StudyNB-302	
	NB-16	NB-32	Placebo	NB-32	Placebo
Number randomized	578	583	581	591	202
Safety Analysis Set	569 (98.4)	573 (98.3)	569 (97.9)	584 (98.8)	200 (99.0)
Efficacy Analysis Sets					
Intention-to-Treat Set	524 (90.7)	538 (90.7)	536 (92.3)	note 1	
Full Analysis Set	471 (81.5)	471 (80.8)	511 (88.0)	482 (81.6)	193 (95.5)
Per Protocol Set	263 (45.5)	267 (45.5)	251 (43.2)	245 (41.5)	92 (45.5)
Completers Analysis Set, Week 56	284 (49.1)	296 (50.8)	290 (49.9)	301 (50.9)[2]	106 (52.5)

	Study NB-303		Study NB-304	
	NB-32	Placebo	NB-32	Placebo
Number randomized	1001	495	335	170
Safety Analysis Sets	992 (99.1)	492 (99.4)	333 (99.4)	169 (99.4)
Efficacy Analysis Sets				
Intention-to-Treat Set	943 (94.2)	474 (95.8)	321 (95.8)	166 (97.6)
Full Analysis Set	825 (82.4)	456 (92.1)	265 (79.1)	159 (93.5)
Per Protocol Set	483 (48.3)	248 (50.1)	149 (44.5)	102 (60.0)
Completers Analysis Set, Week 56	538 (53.7)	267 (53.9)	175 (52.2)	100 (58.8)
Completers Analysis Set, Week 28[3]	619 (61.8)	319 (64.4)	note 3	

Notes

[1] Study NB-302 did not define an intention-to-treat set

[2] A discrepancy in the number per arm in the completers analysis set in Study NB-302 and the total number per arm classified as "completed" in the disposition analysis has been noted by this reviewer and by the applicant. This is attributed to the definition of "completer" that was used for the analysis population. For more information, see Part 3.1.2.3 of this review.

[3] Study NB-303 defined a completers analysis set for week 28

Sources: Study NB301 report, Table 6; Study NB302 report, Table 4; Study NB303 report, Table 4; Study NB304 report, Table 5

3.1.4. Primary efficacy endpoint

For Study NB-301, NB-302 and NB-304, the applicant defined two co-primary efficacy endpoints: (1) the proportion of subjects achieving ≥ 5% reduction in body weight at the end of 56 weeks of treatment and (2) the change from baseline to the end of week 56. These are the co-primary endpoints that are described in the weight management guidance.

Study NB-303 evaluated these endpoints at the end of week 28 as the primary endpoint. This was in response to recommendations from the biometrics review team and the medical division (see letters dated June 9, 2009 and June 23, 2009, submitted under IND 068858). The reason for

this recommendation was the differential treatment of subjects randomized to the NB32 arm and the placebo arm after week 28, as described in the 6/9/09 letter:

> The reason for our disagreement is the differential treatment of the two randomized study arms which begins at week 28 with the re-randomization protocol and continues through week 44. We believe that this differential treatment of the two study arms is likely to introduce bias into the estimate of the efficacy of the efficacy of the naltrexone 36 mg / bupropion 360 mg/day (nal36/bup) at week 56. The estimated placebo-adjusted weight loss attributed to nal36/bup may tend to be larger than it actually is in the target population.
>
> The re-randomization of subjects who have not lost at least 5% of baseline body weight by week 28 initiates a differential treatment of subjects depending on their randomized assignment. A subject in the placebo arm continues on the placebo. A subject in the nal32/bup360 arm is re-randomized, with a 50% chance of being assigned to a higher dose, naltrexone 48 mg/bupropion 360 mg/day (nal48/bup360), and a 50% chance of staying with the nal32/bup360. This re-randomization protocol continues for weeks 32 through 44. This means that a subject who has not yet been assigned to the nal48/bup360 dose may be re-evaluated every 4 weeks from week 28 through week 44, with the possibility of re-randomization to the nal48/bup360 dose at each evaluation, conditional on having been randomized to the nal32/bup360 arm and having not lost at least 5% of baseline body weight.
>
> We believe that the re-randomization protocol is likely to result in a biased overestimate of the efficacy of the nal32/bup360 dose in comparison with the placebo. The basis for this belief is the assumption that subjects in the nal32/bup360 arm who do not lose at least 5% of baseline body weight by week 28 may be non-responders to this dose. They may be similar to patients treated by placebo with regard to the trend in body weight from week 28 to week 56. The potential for a biased overestimate is caused by (1) using the last observation carried forward (LOCF) at the point of assignment to the nal48/bup360 dose for subjects in this subgroup who are in the nal32/bup360 arm, while (2) continuing to include weight records for subjects in this subgroup who are in the placebo arm.
>
> These reasons support our recommendation to use week 28 as the primary efficacy endpoint period. Results at week 56 can be used in supportive analyses.

Based on the 2007 weight management guidance, the efficacy of Contrave would be supported if either one or both the co-primary endpoints were statistically significant (see Part IV-B-3-c). However, the ICH-E9 guidance advises that in the event that a protocol identifies more than one primary endpoint, "the effect on the Type I error should be explained because of the potential for multiplicity problems …; the method of controlling Type I error should be given in the protocol."[10] This places the weight management guidance at variance with the recommendation from the ICH-E9 guidance. The inferential tests for each co-primary endpoint between an active treatment arm and placebo were assessed against a two-sided significance level of 0.05 separately. The protocols state that the tests conducted for each endpoint "must be significant versus the two-sided significance level of 0.05 (i.e. both p-values < 0.05)." This statement can

[10] Part II.B.5., *Guidance for Industry, E9 Statistical Principles for Clinical Trials*, September 1998

be interpreted to mean that the efficacy of Contrave would be supported only if both co-primary endpoints were significant.

3.1.2.4 Statistical analysis methods for primary efficacy endpoint

Primary analysis model: The primary analysis was performed for the FAS analysis set, using last non-missing observation carried forward (LOCF). For percent change from baseline, an analysis of covariance (ANCOVA) model included treatment group and study center as the main effects with baseline measurement as the covariate. The primary analysis for the percentage of subjects achieving ≥ 5% weight loss from baseline was based on a linear logistic regression model, using treatment group and study center as the main effects with the baseline measurement as the covariate. The biostatistics review team concurred with the analysis plan (see letter dated June 9, 2009).

Sensitivity analyses of the co-primary endpoints: In my opinion, the applicant used a reasonable set of analyses that explored the effect of different analysis sets and analysis models on the study conclusions. The sensitivity analysis included plausible interpretations of the weight response of subjects who discontinued. As I discussed in part 3.1 of this review, subjects who withdrew from study medication early had lost less weight at the time of withdrawal, compared to the average weight loss at the same study visit in subjects who completed 56 weeks of study medication. The sensitivity analysis included the following approaches:

a) The primary analysis model with different analysis sets, for both co-primary endpoints. These analysis sets included the ITT analysis set, the completers set and the per protocol set.

b) A repeated measures linear mixed-effects model for the continuous endpoint. This analysis used the ITT analysis set with no estimation for missing data. The analysis model had a random subject effect, fixed class effects for treatment, time (i.e., week), study center, the treatment-by-time interaction, with baseline as a covariate. The Kenward-Rogers approximation was used to estimate the denominator degrees of freedom.

c) A weight regain imputation method, for both co-primary endpoints. The primary analysis model was used, for all randomized subjects, but with an estimate of weight at week 56 in subjects who discontinued early, based on a rate of 0.3 kg of regained weight per month[11]. The estimate was bounded by the subject's baseline weight. If a subject did not return after enrollment (i.e., the subject had no post-baseline weights), the baseline was imputed for all missing values.

d) A baseline carried forward imputation method, for both co-primary endpoints. For all randomized subjects, the endpoint was defined as the week 56 measurement, irrespective of being on study drug or not. For randomized subjects who discontinued active study drug prior to week 56, the endpoint was the baseline measurement (i.e., the percent change from baseline was

[11] See Sacks FM, Bray GA, Carey VJ, Smith SR, Ryan DH, Anton SD et al, 2009. Comparison of weight-loss diets with different compositions of fat, protein and carbohydrates. NEJM 360:859-873.

equal to zero for these subjects). For all subjects who discontinued active study drug prior to week 56, the baseline weight was used to estimate percent change from baseline. This means that subjects who discontinued study drug were estimated to have 0% change from baseline. This imputation was applied to all randomized subjects. The primary analysis models were used with this imputation method.

The baseline carried forward method was applied to week 28 for the primary analysis of Study NB-303.

Adjustment for multiplicity with more than one NB dose: Study NB-301, which had two dose levels of Contrave, conducted the inferential tests associated with the co-primary comparisons in a stepwise manner beginning with the higher active treatment group. The inferential tests for each co-primary endpoint between an active treatment arm and placebo will be assessed against a two-sided significance of 0.05 separately. That is, the comparisons for each co-primary endpoint were each tested at a significance level of 0.05 for N32 vs. placebo only. If each endpoint was significant, $p<0.01$) for this treatment comparison, then the comparison for NB15 was conducted for each co-primary endpoint at a significance level of 0.05.

3.1.2.5 Results of the statistical analysis of efficacy

Continuous endpoint: After 56 weeks of treatment with naltrexone 32 mg / bupropion 360 mg subjects lost a statistically significant amount of weight. The primary endpoint was evaluated at 56 weeks for Study NB-301, NB-302 and NB-304, and at 28 weeks for Study NB-303. The placebo-adjusted mean effect of NB32 was somewhat less than 5%, the benchmark for clinical significance, but the 95% confidence intervals include 5% (TABLE 12, TABLE 14, TABLE 16, TABLE 18).

The majority of subjects who dropped out prior to the end of each study remained within ± 5% of their baseline body weight (FIGURE 6 - FIGURE 9, top portion of each bar). These are the subjects whose final weight was estimated by LOCF in the primary analysis.

Results from sensitivity analyses, including alternative analysis models and different analysis populations, supported the conclusion of the efficacy of NB32 in the continuous endpoint. Results were generally statistically significant, with placebo-adjusted mean effect somewhat less than 5%, and with a 95% CI including 5% (TABLE 12, TABLE 14, TABLE 16, TABLE 18). The placebo-adjusted mean effect of NB32 was generally greater in the completers and per protocol analysis sets than in the baseline-carried-forward and weight-regain analysis sets, as would be expected.

A summary of the results for the continuous weight loss endpoint in each study is as follows:

- Study NB-301 included two dose levels of the combination product NB32 and NB16. The average amount of weight lost in the NB32 arm was greater than the average weight loss in

the NB16 arm (TABLE 12). This result supports a dose-response relationship between these two dosages.

- Study NB-302 was conducted with an intensive program for behavior modification in both arms. The effectiveness of this program is demonstrated by the larger mean change from baseline in both the NB32 and placebo arms, compared with Study NB-301 (TABLE 14). However, the placebo-adjusted effect of NB32 in Study NB-302 is somewhat smaller than the effect in Study NB-301. This finding is consistent across the primary analysis and supportive analyses. Subjects in the placebo group of NB-302, with the intensive behavior modification program, may have been more successful in losing weight than subjects in the placebo group of NB-301.

- Study NB-303 evaluated the primary endpoint at week 28. However, the effect of NB28 at week 28 may be fairly similar to the effect at week 56. The longitudinal profile of mean weight loss in completers appears to stabilize at week 28 in all four studies (FIGURE 2 - FIGURE 5).

- Study NB-304 was conducted in obese subjects with type 2 diabetes. The placebo-adjusted effect of NB32 at week 56 was smaller than the effect in the other three studies (TABLE 18). This finding is consistent with the clinical expectation for less weight loss in diabetic subjects compared with non-diabetic subjects.

Categorical endpoint: After one year of treatment with NB32 (or 28 weeks in the case of Study NB-303), a statistically significantly greater percentage lost at least 5% of their baseline body weight, compared to placebo (TABLE 13, TABLE 15, TABLE 17, TABLE 19). The results from the analysis of the FAS were supported by the results from sensitivity analyses using other versions of the analysis data sets. This result supports the criterion for statistical significance in the 5% responder endpoint, as described in the weight management guidance.

Of the several sensitivity analyses, the analysis using the baseline observation carried forward (BOCF) may be of greatest interest. This is because: (1) the BOCF analysis can be applied to all randomized subjects; and (2) it may be reasonable to assume that most subjects who drop out of a weight loss study can be classified as non-responders. However, subjects also dropped out of the NB studies who were 5% responders at the time of dropout. For example, in Study NB-301, 226/471 (48%) of NB32 subjects were responders in the FAS/LOCF analysis set, while 180/583 (32%) were responders in the BOCF analysis set. The differences reflect 112 subjects who were randomized but were not in the FAS (having dropped out before providing any post-baseline weights), and 46 subjects who were responders at the time they dropped out. A similar finding occurs in the other 3 studies as well. Some subjects in the NB32 arm dropped out late in a study as responders (see the later weeks in FIGURE 2 - FIGURE 4). One interpretation of this pattern is that late dropout-responders in the NB32 arm were satisfied with their weight loss and no longer willing to tolerate the adverse events associated with NB32.

TABLE 12 Study NB-301; Body weight (kg), percent change from baseline to week 56 endpoint; Primary analysis and sensitivity analyses

Study NB-301: Customary diet and behavioral counseling	Group	n	Baseline mean (SD)	Percent Change from Baseline LSMean (SE)	LS Mean Difference from Placebo (95% CI)	p
1A. Full Analysis Set (LOCF): reported by applicant	Placebo	511	99.3 (14.3)	-1.3 (0.3)		
	NB16	471	100.1 (14.4)	-5.9 (0.3)	-3.7 (-4.5, -2.9)	<0.001
	NB32	471	100.2 (16.3)	-6.1 (0.3)	-4.8 (-5.6, -4.0)	<0.001
1B. Full Analysis Set (LOCF): this reviewer's analysis	Placebo	511	99.3 (14.3)	-1.4 (0.3)		
	NB16	471	100.1 (14.4)	-4.9 (0.3)	-3.5 (-4.3, -2.7)	<0.001
	NB32	471	100.2 (16.3)	-6.0 (0.3)	-4.6 (-5.4, -3.8)	<0.001
2. Completers Analysis Set	Placebo	290	99.2 (14.6)	-1.8 (0.5)		
	NB16	284	99.8 (14.6)	-6.7 (0.5)	-4.9 (-6.1, -3.6)	<0.001
	NB32	296	99.8 (16.4)	-8.1 (0.5)	-6.2 (-7.5, -5.0)	<0.001
3. ITT Analysis Set	Placebo	536	99.5 (14.4)	-1.3 (0.3)		
	NB16	524	99.5 (14.5)	-4.5 (0.3)	-3.2 (-4.0, -2.4)	<0.001
	NB32	538	99.8 (16.1)	-5.4 (0.3)	-4.1 (-4.9, -3.3)	<0.001
4. Weight Regain Imputation Method	Placebo	581	99.5 (14.3)	-1.2 (0.3)		
	NB16	578	99.5 (14.8)	-3.7 (0.3)	-2.5 (-3.2, -1.8)	<0.001
	NB32	583	99.7 (14.8)	-4.6 (0.3)	-3.4 (-4.1, -2.7)	<0.001
5. Baseline Carried Forward Analysis	Placebo	581	99.5 (14.3)	-0.9 (0.3)		
	NB16	578	99.5 (14.3)	-3.3 (0.3)	-2.4 (-3.1, -1.7)	<0.001
	NB32	583	99.7 (15.9)	-4.0 (0.3)	-3.1 (-3.8, -2.4)	<0.001
6. Mixed Model Repeated Measures (FAS)	Placebo	511	99.2 (14.5)	-1.0 (0.4)		
	NB16	471	99.5 (14.4)	-5.2 (0.4)	-4.3 (-5.3, -3.2)	<0.001
	NB32	471	99.9 (16.3)	-6.6 (0.4)	-5.7 (-6.7, -4.6)	<0.001
7. Per Protocol Set	Placebo	251	99.3 (14.9)	-2.3 (0.5)		
	NB16	262	100.6 (14.9)	-7.1 (0.5)	-4.7 (-6.1, -3.4)	<0.001
	NB32	267	100.3 (17.0)	-8.3 (0.5)	-6.0 (-7.3, -4.6)	<0.001

Source: Table ISE.301.1-6, Table 14.2-5

TABLE 13 Study NB-301; Body weight, proportion 5% responders at week 56; primary analysis and sensitivity analyses.

Study NB-301: Customary diet and behavioral counseling	Group	n	5% responders n(%)	Odds Ratio (vs. placebo)	95% CI of Odds Ratio	p
1A. Full Analysis Set (LOCF)	Placebo	511	84 (16.4%)			
	NB16	471	186 (39.5%)	3.4	(2.5, 4.6)	<0.001
	NB32	471	226 (48.0%)	4.9	(3.6, 6.6)	<0.001
1B. Full Analysis Set (LOCF): this reviewer's analysis	Placebo	511	84 (16.4%)			
	NB16	471	186 (39.5%)	3.4	(2.5, 4.6)	<0.001
	NB32	471	226 (48.0%)	4.9	(3.6, 6.6)	<0.001
2. Completers Analysis Set	Placebo	290	67 (23.1%)			
	NB16	284	155 (54.6%)	4.2	(2.9, 6.1)	<0.001
	NB32	296	183 (61.8%)	5.8	(4.0, 8.3)	<0.001
3. ITT Analysis Set	Placebo	536	93 (17.4%)			
	NB16	524	190 (36.3%)	2.8	(2.1, 3.7)	<0.001
	NB32	538	226 (42.0%)	3.6	(2.7, 4.8)	<0.001
4. Weight Regain Imputation Method	Placebo	581	78 (13.4%)			
	NB16	578	175 (30.3%)	2.9	(2.1, 3.9)	<0.001
	NB32	583	203 (34.8%)	3.6	(2.7, 4.9)	<0.001
5. Baseline Carried Forward Analysis	Placebo	581	67 (11.5%)			
	NB16	578	156 (27.0%)	2.9	(2.1, 4.0)	<0.001
	NB32	583	180 (30.9%)	3.6	(2.6, 4.9)	<0.001
6. Per Protocol Set	Placebo	251	67 (26.7%)			
	NB16	263	141 (52.6%)	3.4	(2.3, 5.0)	<0.001
	NB32	267	162 (60.7%)	4.6	(3.1, 6.8)	<0.001

Source: Table ISE.301.1-7, Table 14.2-20

TABLE 14 Study NB-302; Body weight (kg), percent change from baseline to week 56 endpoint; Primary analysis and sensitivity analyses

Study NB-302: Intensive program for behavior modification	Group	n	Baseline mean (SD)	Percent Change from Baseline LSMean (SE)	LS Mean Difference from Placebo (95% CI)	p
1. Full Analysis Set (LOCF)[1]	Placebo	193	101.9 (15.0)	-5.1 (0.6)		
	NB32	482	100.7 (15.4)	-9.3 (0.4)	-4.2 (-5.6, -2.9)	<0.001
2. Completers Analysis Set	Placebo	106	100.4 (14.3)	-7.3 (0.9)		
	NB32	301	101.2 (15.1)	-11.5 (0.6)	-4.2 (-6.1, -2.4)	<0.001
3. ITT Analysis Set	Placebo	196	101.8 (15.0)	-4.9 (0.6)		
	NB32	565	100.3 (15.5)	-8.1 (0.4)	-3.2 (-4.5, -1.8)	<0.001
4. Weight Regain Imputation Method	Placebo	202	101.9 (15.0)	-4.9 (0.6)		
	NB32	591	100.2 (15.4)	-7.3 (0.4)	-2.4 (-3.7, -1.1)	<0.001
5. Baseline Carried Forward Analysis	Placebo	202	101.9 (15.0)	-4.2 (0.6)		
	NB32	591	100.2 (15.4)	-6.4 (0.4)	-2.2 (-3.5, -0.9)	<0.001
6. Mixed Model Repeated Measures (FAS)	Placebo	193	100.4 (14.3)	-5.4 (0.7)		
	NB32	482	101.2 (15.1)	-10.3 (0.4)	-5.0 (-6.6, -3.3)	<0.001
7. Per Protocol	Placebo	92	101.3 (14.7)	-8.0 (1.0)		
	NB32	245	99.3 (14.6)	-12.0 (0.7)	-4.0 (-6.1, -1.9)	<0.001

Note 1: I confirmed the results for the FAS.

Source: Table ISE.302.1-6, Table 14.2-5

TABLE 15 Study NB-302; Body weight, proportion 5% responders at week 56; primary analysis and sensitivity analyses.

Study NB-302: Intensive program for behavior modification	Group	n	5% responders n(%)	Odds Ratio (vs. placebo)	95% CI of Odds Ratio	p
1. Full Analysis Set (LOCF) [1]	Placebo	193	82 (42.5%)			
	NB32	482	320 (66.4%)	2.8	(2.0, 4.1)	<0.001
2. Completers Analysis Set	Placebo	106	64 (60.4%)			
	NB32	301	242 (80.4%)	2.9	(1.8, 4.8)	<0.001
3. ITT Analysis Set	Placebo	196	84 (42.9%)			
	NB32	565	321 (56.8%)	1.8	(1.3, 2.6)	<0.001
4. Weight Regain Imputation Method	Placebo	202	77 (38.1%)			
	NB32	591	304 (51.4%)	1.8	(1.3, 2.5)	<0.001
5. Baseline Carried Forward Analysis	Placebo	202	68 (33.7%)			
	NB32	591	269 (45.5%)	1.7	(1.2, 2.4)	<0.001
6. Per Protocol Set	Placebo	92	57 (62.0%)			
	NB32	145	198 (80.8%)	3.0	(1.7, 5.2)	<0.001

Note 1: I confirmed the results for the FAS.

Source: Table ISE.302.1-7, Table 14.2-19

TABLE 16 Study NB-303; Body weight (kg), percent change from baseline to week 28 endpoint; Primary analysis and sensitivity analyses

Study NB-303: Customary diet and behavioral counseling	Group	n	Baseline mean (SD)	Percent Change from Baseline LSMean (SE)	LS Mean Difference from Placebo (95% CI)	p
1. Full Analysis Set (LOCF) [1]	Placebo	456	99.3 (16.0)	-1.9 (0.3)		
	NB32	825	100.7 (16.7)	-6.5 (0.2)	-4.6 (-5.2, -3.9)	<0.001
2. Completers Analysis Set	Placebo	319	99.0 (15.99)	-2.4 (0.3)		
	NB32	619	101.2 (17.1)	-7.8 (0.2)	-5.4 (-6.2, -4.6)	<0.001
3. ITT Analysis Set	Placebo	474	99.4 (15.9)	-1.9 (0.3)		
	NB32	943	100.4 (16.7)	-5.7 (0.2)	-3.9 (-4.5, -3.2)	<0.001
4. Weight Regain Imputation Method	Placebo	495	99.2 (15.9)	-1.9 (0.3)		
	NB32	1001	100.3(16.6)	-5.2 (0.2)	-3.4 (-3.9, -2.8)	<0.001
5. Baseline Carried Forward Analysis	Placebo	495	99.2(15.9)	-1.5 (0.3)		
	NB32	1001	100.3 (16.6)	-4.8 (0.2)	-3.3 (-3.9, -2.7)	<0.001
6. Mixed Model Repeated Measures (FAS)	Placebo	456	99.1 (15.9)	-2.1 (0.3)		
	NB32	825	101.0 (17.0)	-7.2 (0.2)	-5.1 (-5.8, -4.4)	<0.001
7. Per Protocol Set	Placebo	248	98.5 (15.9)	-2.8 (0.4)		
	NB32	483	101.2 (17.2)	-8.2 (0.3)	-5.4 (-6.3, -4.5)	<0.001

Note 1: I confirmed the results for the FAS.

Source: Table ISE.303.1-6A, Table 14.2-9

TABLE 17 Study NB-303 at week 28; Body weight, proportion 5% responders at week 28; primary analysis and sensitivity analyses.

Study NB-303: Customary diet and behavioral counseling	Group	n	5% responders n(%)	Odds Ratio (vs. placebo)	95% CI of Odds Ratio	p
1. Full Analysis Set (LOCF) [1]	Placebo	456	80 (17.5%)			
	NB32	825	459 (55.6%)	6.6	(5.0, 8.8)	<0.001
2. Completers Analysis Set	Placebo	319	71 (22.3%)			
	NB32	619	426 (68.8%)	8.7	(6.2, 12.1)	<0.001
3. ITT Analysis Set	Placebo	474	81 (17.1%)			
	NB32	943	461 (48.9%)	5.1	(3.9, 6.8)	<0.001
4. Weight Regain Imputation Method	Placebo	495	79 (16.0%)			
	NB32	1001	446 (44.6%)	4.7	(3.5, 6.2)	<0.001
5. Baseline Carried Forward Analysis	Placebo	495	69 (13.9%)			
	NB32	1001	421 (42.1%)	4.9	(3.7, 6.6)	<0.001
6. Per Protocol Set	Placebo	248	64 (25.8%)			
	NB32	483	345 (71.4%)	8.4	(5.7, 12.2)	<0.001

Note 1: I confirmed the results for the FAS.

Source: Table ISE.303.1-7A, Table 14.2-37

TABLE 18 Study NB-304; Body weight (kg), percent change from baseline to week 56 endpoint; Primary analysis and sensitivity analyses

Study NB-304: Obese subjects with type 2 diabetes	Group	n	Baseline mean (SD)	Percent Change from Baseline LSMean (SE)	LS Mean Difference from Placebo (95% CI)	p
1A. Full Analysis Set (LOCF)	Placebo	159	105.0 (17.1)	-1.8 (0.4)		
	NB32	265	106.5 (19.1)	-5.0 (0.3)	-3.3 (-4.3, -2.2)	<0.001
1B. Full Analysis Set (LOCF): this reviewer's analysis	Placebo	159	105.0 (17.1)	-2.1 (0.5)		
	NB32	265	106.4 (19.1)	-5.5 (0.4)	-3.4 (-4.5, -2.3)	<0.001
2. Completers Analysis Set	Placebo	100	105.1 (16.9)	-2.2 (0.6)		
	NB32	175	107.0 (19.5)	-5.9 (0.5)	-3.7 (-5.2, -2.2)	<0.001
3. ITT Analysis Set	Placebo	166	105.3 (16.9)	-1.7 (0.4)		
	NB32	321	104.2 (19.1)	-3.7 (0.3)	-2.0 (-3.0, -1.0)	<0.001
4. Weight Regain Imputation Method	Placebo	170	105.1 (17.0)	-1.7 (0.4)		
	NB32	335	104.2 (18.9)	-3.5 (0.3)	-1.9 (-2.8, -0.9)	<0.001
5. Baseline Carried Forward Analysis	Placebo	170	105.1 (17.0)	-1.3 (0.4)		
	NB32	335	104.2 (18.9)	-3.1 (0.3)	-1.7 (-2.7, -0.8)	<0.001
6. Mixed Model Repeated Measures (FAS)	Placebo	159	105.3 (17.1)	-1.9 (0.5)		
	NB32	265	107.1 (19.3)	-5.6 (0.4)	-3.7 (-5.0, -2.4)	<0.001
7. Per Protocol Set	Placebo	102	104.4 (17.4)	-2.0 (0.6)		
	NB32	149	107.7 (20.1)	-6.1 (0.5)	-4.2 (-5.7, -2.6)	<0.001

Source: Table ISE.304.1-6, Table 14.2-5

TABLE 19 Study NB-304; Body weight, proportion 5% responders at week 56; primary analysis and sensitivity analyses.

Study NB-304: Obese subjects with type 2 diabetes	Group	n	5% responders n(%)	Odds Ratio (vs. placebo)	95% CI of Odds Ratio	p
Full Analysis Set (LOCF) [1]	Placebo	159	30 (18.9%)			
	NB32	265	118 (44.5%)	3.4	(2.2, 5.5)	<0.001
Completers Analysis Set	Placebo	100	24 (24.0%)			
	NB32	175	93 (53.1%)	3.7	(2.1, 6.5)	<0.001
ITT Analysis Set	Placebo	166	30 (18.1%)			
	NB32	321	115 (35.8%)	2.5	(1.6, 4.0)	<0.001
Weight Regain Imputation Method	Placebo	170	27 (15.9%)			
	NB32	335	104 (31.0%)	2.4	(1.5, 3.8)	<0.001
Baseline Carried Forward Analysis	Placebo	170	24 (14.1%)			
	NB32	335	94 (28.1%)	2.4	(1.4, 3.9)	<0.001
Per Protocol Set	Placebo	102	25 (24.5%)			
	NB32	149	82 (55.0%)	4.0	(2.3, 7.1)	<0.001

Note 1: I confirmed the results for the FAS, after removing the term "site" from the model (based on problems with model validity with the term "site" included).

Source: Table ISE.304.1-7, Table 14.2-22

FIGURE 6 Study NB-301; Distribution of weight change at week 52; FAS analysis set

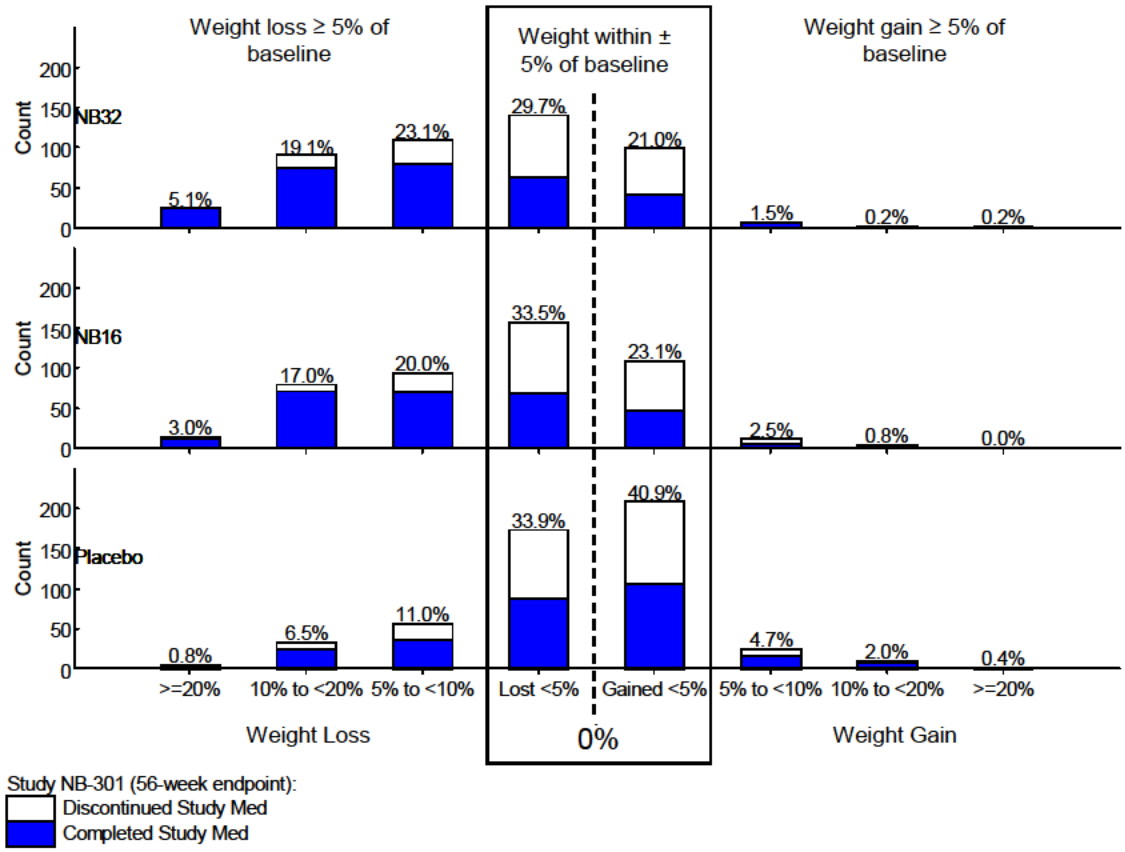


FIGURE 7 Study NB-302; Distribution of weight change at week 56; FAS analysis set

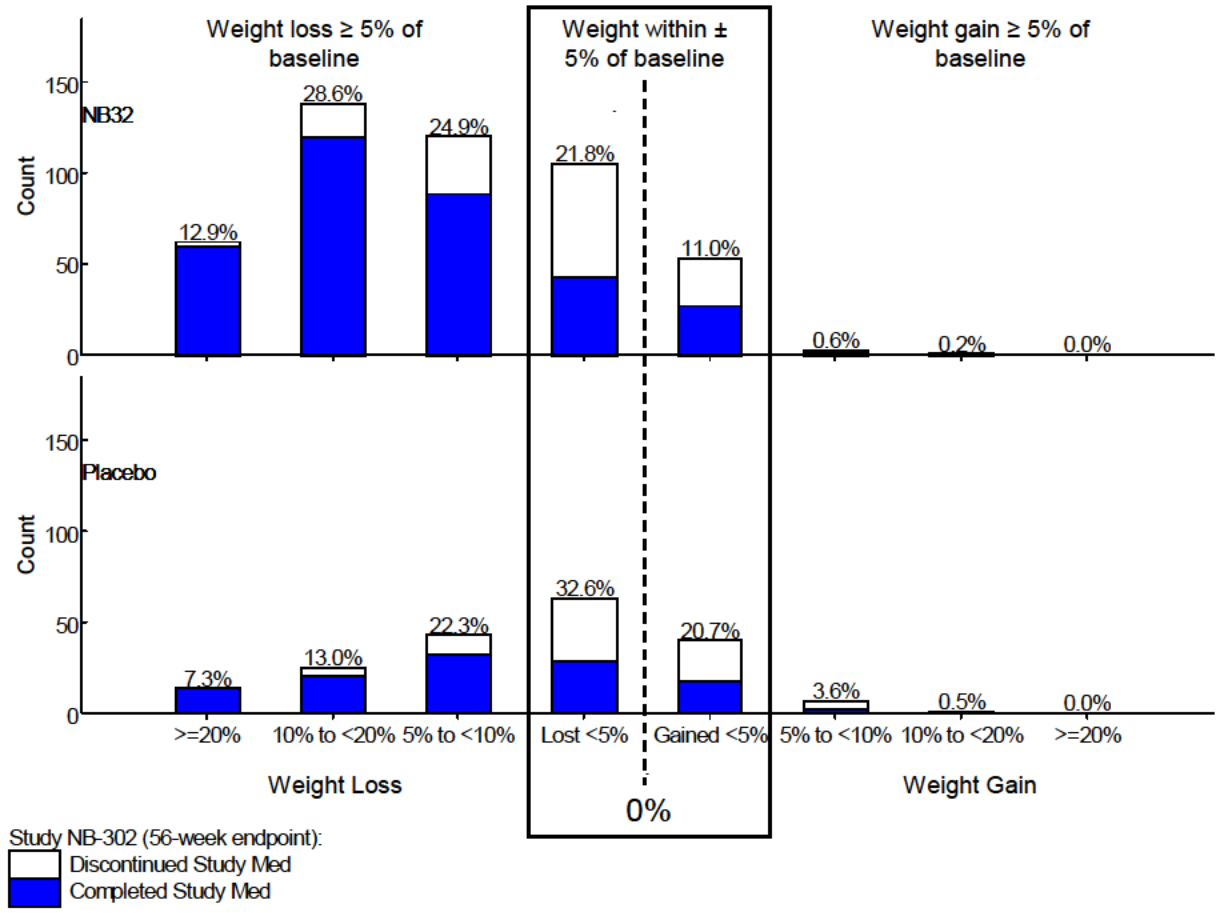


FIGURE 8 Study NB-303; Distribution of weight change at week 28; FAS analysis set

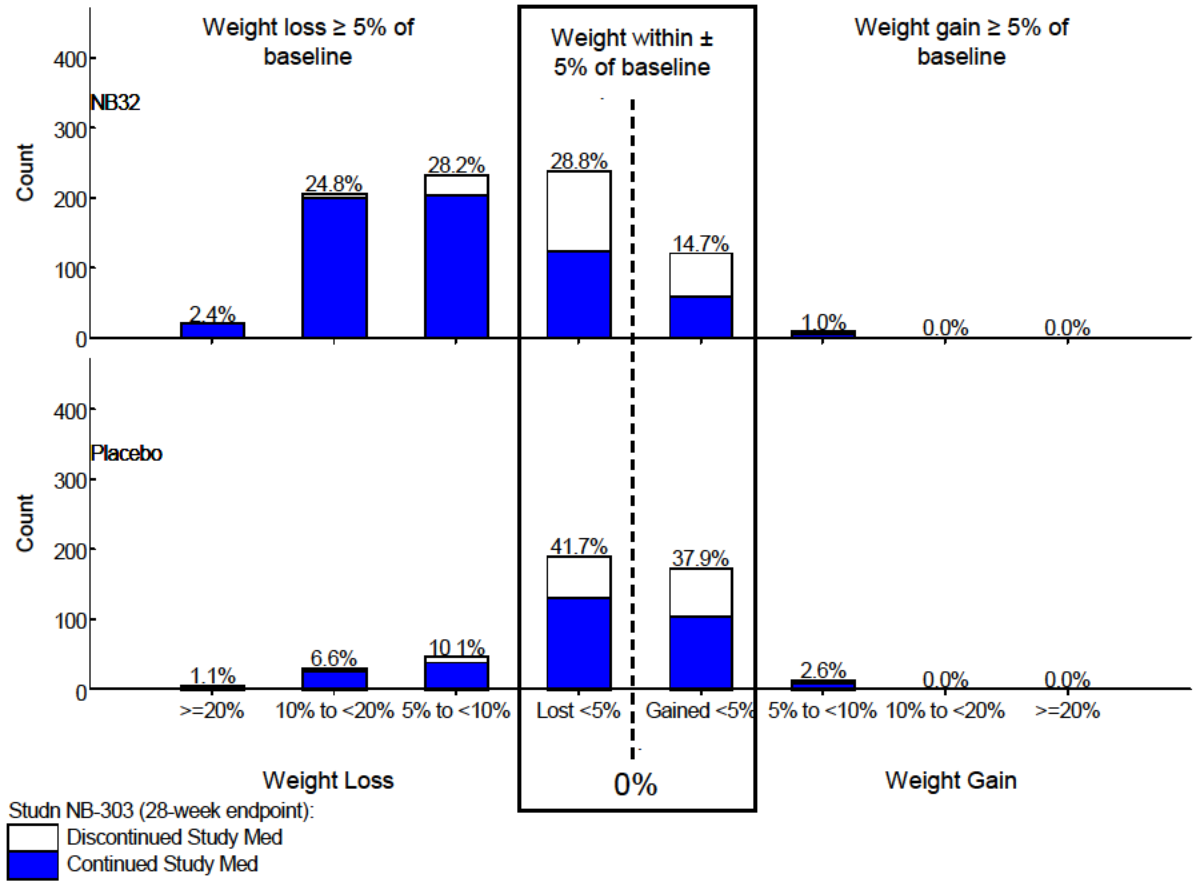
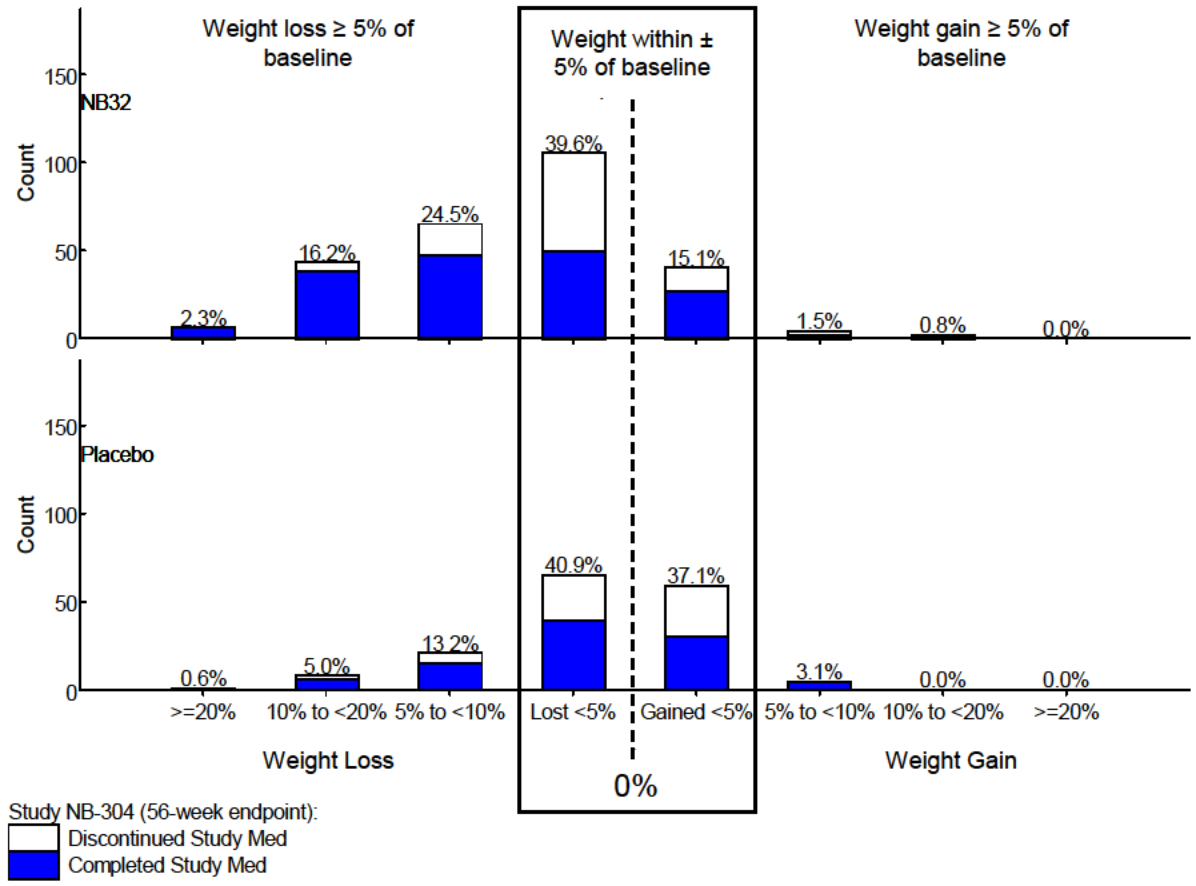


FIGURE 9 Study NB-304; Distribution of weight change at week 56; FAS analysis set



3.1.2.6 Key secondary efficacy endpoints

The protocol of each study described a set of 15-20 key secondary efficacy endpoints (TABLE 20). These endpoints were only evaluated if both of the co-primary efficacy endpoints were statistically significant. This condition was met in all four Phase 3 studies. The testing of the key secondary efficacy endpoints was structured in a sequential manner, using a (b) (4) approach to control for Type I error. Testing of the secondary endpoints continued in the designated sequence until the first instance of a p-value > 0.05. At that point, results from statistical comparisons of NB to placebo in the remaining sequence of endpoints were not considered. For Study NB-301, which had two dose levels of NB, this approach was carried out separately for NB32 and for NB16. For purposes of clarity, I focus on the sequential testing of NB32 compared to placebo in each study.

Fourteen of the key secondary efficacy endpoints were the same in all four studies (TABLE 20A). An additional endpoint, relating to the Control of Eating Questionnaire, was modified from the total score in Study NB-302 to the single item #19 in the other studies. Because Study NB-302 was finished approximately a year before the other three studies were finished, the refinement of this endpoint may be related to the results from Study NB-302. Study NB-303 included three additional endpoints that related to the evaluation of weight change at week 56 because the primary endpoint period for this study was week 28 (TABLE 20B). Study NB-304 included five additional endpoints that related to glycemic control, because this study enrolled obese subjects with Type 2 diabetes (TABLE 20C).

Among the fifteen key secondary endpoints in common to all four studies, the relative testing order was modified in each study (TABLE 20A). A summary of the results of the closed testing procedure in each study is as follows:

- In Study NB-301, the testing sequence stopped at the 11th endpoint in Study NB-301 (fasting LDL; p=0.484). None of the remaining four endpoints had nominal p-values < 0.05 in a direction that supported the efficacy of NB-32. Two of these later endpoints were systolic and diastolic blood pressure. These did have nominal p-values < 0.05 but the effect of NB-32 was to increase SDP and DBP compared to placebo, which does not support the efficacy of NB-32 (TABLE 28, TABLE 29).

- In Study NB-302, the testing sequence stopped at the 8th endpoint (hs-CRP; p=0.165). None of the remaining 7 endpoints had nominal p-values < 0.05 in a direction that supported the efficacy of NB-32. As with Study NB-301, two of the later endpoints were SDP and DBP, which had nominal p-values < 0.05 in the direction of inferiority of NB-32 to placebo ((TABLE 28, TABLE 29).

- In Study NB-303, the testing sequence stopped at the 8th endpoint (hs-CRP; p=0.091) and 4 of the remaining 9 endpoints had nominal p-values < 0.05 in a direction that supported the efficacy of NB-32. One of the later endpoints, DBP, had a nominal p-value < 0.05 in the direction of inferiority of NB-32 to placebo (TABLE 29).

- In Study NB-304, the testing sequence stopped at the 4th endpoint (fasting blood glucose level; p=0.065), and 6 of the remaining 16 endpoints had nominal p-values < 0.05 in a direction that supported the efficacy of NB-32. The results for HbA1c, the first endpoint in the sequence, support the improved glycemic control in the NB-32 arm compared to the placebo arm in the diabetic patients in this study (TABLE 27).

Considering the three studies conducted in non-diabetic subjects, five endpoints consistently had a statistically significant placebo-adjusted effect of NB32 and were also considered as part of the sequence of tests (TABLE 20A):

- the proportion of subjects with ≥ 10% decrease in body weight
- waist circumference
- fasting HDL level
- fasting triglycerides level
- Impact of Weight on Quality of Life (IWQOL-Lite) total score

The results of these endpoints support the efficacy of NB-32 and NB-16 compared to placebo (TABLE 21 - TABLE 26). Two of these endpoints, fasting triglyceride level and fasting HDL level, were identified as significant effects in the testing sequences for all four studies. From a statistical perspective, these findings could be included in the package insert. A final decision about labeling depends on an assessment of their clinical significance.

The applicant analyzed triglycerides as a log transform, with the results back-transformed as a percentage change from baseline (TABLE 24). This approach has the benefit of stabilizing the variance and reducing the influence of outliers, and may be a reasonable approach to the analysis of triglyceride data, which can be skewed with outlying high values. However, a drawback to this approach is its interpretability from the clinical perspective. Clinicians are more accustomed to evaluating the change from baseline in triglycerides on the original scale of measurement. For this reason, I also analyzed triglycerides on the original scale of measurement. Results from Study NB-303 (week 28) were different on the two scales, with a non-significant comparison between NB32 and placebo on the original scale of measurement (TABLE 25). This would have stopped the sequence of testing at step 6 rather than at step 8, excluding both triglycerides and IWQOL-Lite total score from the set of significant results (TABLE 20). Results from the analysis of triglycerides on the original scale of measurement would not have changed the testing sequence in the other studies (TABLE 24, TABLE 25).

Although this approach to controlling Type I error in secondary endpoints is acceptable from an operational perspective, and the Division and the Biometrics team did not raise objections to it at the protocol stage, the results from this set of clinical studies illustrate a significant limitation, in my opinion. The long list of key secondary endpoints, the differences among studies in the testing order, and the inclusion of additional endpoints in some studies but not others, has reduced the clarity of interpreting the significance of endpoints when viewed across studies. For example, it may be challenging to understand why the efficacy of Contrave is supported for improving the quality of life (IWQOL-Lite) but not for improving insulin resistance (HOMA-IR) in the three studies conducted in non-diabetic subjects. Both endpoints have p-values < 0.05 in all three studies (TABLE 20A). However, the IWQOL-Lite endpoint was included as the 5th, 6th,

and 7th endpoint in studies NB-301, NB-302 and NB-303, respectively, and HOMA-IR was included as the 9th, 7th and 11th endpoint. Because the testing sequence in NB-303 was stopped at the 8th endpoint (with the hsCRP endpoint p-value of 0.091), the p-value of HOMA-IR is not considered in Study NB-303.

I believe that an improved approach to the control of Type I error in a list of secondary endpoints in a future application would impose more structure and consistency among studies, as follows: (1) identify groups of endpoints that are clinically related, such as the serum cholesterol endpoints; (2) within each clinical group, specify a testing order that is consistent from study to study; (3) use a method that allows for this structure and protects the overall Type I error among and within the clinical groups. With this more structured approach, it would be possible to add or remove clinical groups of endpoints from study to study, while still maintaining clarity in interpreting results among studies.

TABLE 20 Key secondary efficacy endpoints; results of analysis, showing pre-specified order for each Phase 3 study

Type of Endpoint: Endpoint Description	**NB-301** Customary diet and behavioral counseling **Test Order**	**p-value, Status of Testing**	**NB-302** Intensive program for behavior modification **Test Order**	**p-value, Status of Testing**	**NB-303** Customary diet and behavioral counseling **Test Order**	**p-value, Status of Testing**	**NB-304** Obese subjects with type 2 diabetes **Test Order**	**p-value, Status of Testing**
A. The following endpoints refer to the change from baseline and the week 56 visit (week 28 for Study NB-303):								
Weight: Proportion of subjects with ≥ 10% decrease in total body weight (10% responders)	**1**	p<0.001 Continue	**1**	p<0.001 Continue	**3**	p<0.001 Continue	6	Stopped (p<0.001)
Body measurement: Waist circumference	**2**	p<0.001 Continue	**2**	p<0.001 Continue	**4**	p<0.001 Continue	5	Stopped (p=0.006)
Lipids: Fasting triglycerides level	**4**	p<0.001 Continue	**3**	p=0.004 Continue	**6**	p=0.007 Continue	**2**	p=0.007 Continue
Glucose/Insulin: Fasting insulin level	**7**	p<0.001 Continue	**4**	p=0.003 Continue	9	Stopped (p<0.001)	11	Stopped (p=0.563)
Lipids: Fasting HDL level	**3**	p<0.001 Continue	**5**	p<0.001 Continue	**5**	p<0.001 Continue	**3**	p<0.001 Continue
Patient Reported Outcome: Impact of Weight on Quality of Life (IWQOL-Lite) total score	**5**	p<0.001 Continue	**6**	p=0.001 Continue	**7**	p<0.001 Continue	13	Stopped (p=0.208)
Glucose/Insulin: Homeostatis model assessment – insulin resistance (HOMA-IR) ratio	**9**	p<0.001 Continue	**7**	p=0.003 Continue	11	Stopped (p<0.001)	10	Stopped (p=0.361)
C Reactive Protein: High-sensitivity C reactive protein levels (hs-CRP)	**6**	p=0.008 Continue	**8**	**p=0.165 Stop**	**8**	**p=0.091 Stop**	14	Stopped (p=0.312)
Glucose/Insulin: Fasting blood glucose level	**8**	p=0.010 Continue	9	Stopped (p=0.225)	10	Stopped (p=0.544)	**4**	**p=0.065 Stop**

Type of Endpoint: Endpoint Description	NB-301 Customary diet and behavioral counseling Test Order	NB-301 p-value, Status of Testing	NB-302 Intensive program for behavior modification Test Order	NB-302 p-value, Status of Testing	NB-303 Customary diet and behavioral counseling Test Order	NB-303 p-value, Status of Testing	NB-304 Obese subjects with type 2 diabetes Test Order	NB-304 p-value, Status of Testing
Lipids: Fasting LDL	**11**	**p=0.484 Stop**	10	Stopped (p=0.245)	13	Stopped (p<0.004)	17	Stopped (p=0.641)
Blood Pressure: Systolic blood pressure	12	Stopped (p<0.001) NB inferior to placebo	11	Stopped (p=0.002) NB inferior to placebo	14	Stopped (p=0.556)	18	Stopped (p=0.297)
Blood Pressure: Diastolic blood pressure	13	Stopped (p=0.022) NB inferior to placebo	12	Stopped (p=0.017) NB inferior to placebo	15	Stopped (p=0.017) NB inferior to placebo	19	Stopped (p=0.582)
Patient Reported Outcome: Inventory of Depressive Symptomatology Subject-Rated (IDS-SR) total score	14	Stopped (p=0.102)	13	Stopped (p=0.827)	16	Stopped (p=0.844)	20	Stopped (p=0.002)
Patient Reported Outcome: Food Craving Inventory (FCI) sweets and carbohydrate subscale scores	15	Stopped (p=0.598)	14	Stopped (p=0.776)	17	Stopped (p=0.934)	21	Stopped (p=0.322)
Patient Reported Outcome: Control of Eating Questionnaire – total score	---	---	15	Stopped (p=0.078)	---	---	---	---
Patient Reported Outcome: Control of Eating Questionnaire, Item #19	**10**	p<0.001 Continue	---	---	12	Stopped (p<0.001)	16	Stopped (p=0.090)
B. The following endpoints are specific to Study NB-303								
Weight: Percentage change in total body weight between baseline and week 56					**1**	p<0.001 Continue		
Weight: Proportion of subjects with ≥ 5% decrease in total weight at week 56					**2**	p<0.001 Continue		

Type of Endpoint: Endpoint Description	NB-301 Customary diet and behavioral counseling Test Order	NB-301 p-value, Status of Testing	NB-302 Intensive program for behavior modification Test Order	NB-302 p-value, Status of Testing	NB-303 Customary diet and behavioral counseling Test Order	NB-303 p-value, Status of Testing	NB-304 Obese subjects with type 2 diabetes Test Order	NB-304 p-value, Status of Testing
C. The following endpoints are specific to Study NB-304, and refer to the week 56 visit:								
Glycemic control: HbA1c (change from baseline)							**1**	p<0.001 Continue
Glycemic control: Proportion of subjects with HbA1c < 7%							7	Stopped (p<0.001)
Glycemic control: Percent of subjects requiring rescue medications for diabetes							8	Stopped (p=0.008)
Glycemic control: Percent of subjects requiring change in dose(s) or oral hypoglycemic medication							9	Stopped (p=0.672)
Glycemic control: Proportion of subjects with HbA1c < 6.5%							12	Stopped (p=0.004)
Glycemic control: Percent of subjects discontinuing due to poor glycemic control							15	Stopped
D. Key study dates:								
End of study (Week 56 visit of final subject)		5/26/09		9/12/08		6/8/09		6/1/09
Final protocol amendment		3/30/09		1/9/08		3/31/09		3/30/09
Statistical analysis plan (SAP)		6/15/09		3/24/08		8/25/09		6/15/09

Notes

- Only the comparisons of NB-32 vs placebo are depicted in this table. The comparisons of NB-16 vs placebo are not included.
- p-values shown in parentheses are nominal p-values that are not considered because the testing sequence has stopped at an earlier point in the sequence.
- Analyses are based on the FAS/LOCF analysis population.
- The endpoints are ordered by the test order of Study NB-302 because this study was completed prior to the other studies.

Sources

Study NB-301: Clinical Study Report, Table 14
Study NB-302: Clinical Study Report, Table 12
Study NB-303: Clinical Study Report, Table 12
Study NB-304: Clinical Study Report, Table 15

TABLE 21 Phase 3 studies; Body weight, proportion of 10% responders at week 56 (week 28 for Study NB-303); Full Analysis Set (LOCF)

	Group	n	10% responders n(%)	Odds Ratio (vs. placebo)	95% CI of Odds Ratio	p
Study NB-301 (week 56)	Placebo	511	38 (7.4%)			
Customary diet and	NB16	471	95 (20.2%)	3.2	(2.1, 4.8)	<0.001
behavior counseling	NB32	471	116 (24.6%)	4.2	(2.8, 6.2)	<0.001
Study NB-302 (week 56)	Placebo	193	39 (20.2%)			
Intensive lifestyle modification counseling	NB32	482	200 (41.5%)	2.9	(2.0, 4.4)	<0.001
Study NB-303 (week 28)	Placebo	456	32 (7.0%)			
Customary diet and behavior counseling	NB32	825	225 (27.3%)	5.4	(3.6, 8.0)	<0.001
Study NB-304 (week 56)	Placebo	159	9 (5.7%)			
Obese subjects with type 2 diabetes	NB32	265	49 (18.5%)	3.8	(1.8, 7.9)	<0.001

Sources:
Study NB-301: Clinical Study Report, Table 15
Study NB-302: Clinical Study Report, Table 13
Study NB-303: Clinical Study Report, Table 17
Study NB-304: Clinical Study Report, Table 14.2-47A

TABLE 22 Phase 3 studies: Waist circumference (cm), change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF)

	Group	n	Baseline mean (SD)	Percent Change from Baseline LSMean (SE)	LS Mean Difference from Placebo (95% CI)	p
Study NB-301 (week 56)	Placebo	348	110.1 (12.2)	-2.5 (0.4)		
Customary diet and	NB16	342	109.8 (11.2)	-5.0 (0.4)	-3.7 (-4.5, -2.9)	<0.001
behavior counseling	NB32	356	108.8 (11.3)	-6.2 (0.4)	-4.8 (-5.6, -4.0)	<0.001
Study NB-302 (week 56)	Placebo	141	109.0 (11.8)	-6.8 (0.8)		
Intensive lifestyle modification counseling	NB32	391	109.3 (11.4)	-10.0 (0.5)	-3.2 (-4,8, -1.6)	<0.001
Study NB-303 (week 28)	Placebo	315	108.9 (11.7)	-2.7 (0.4)		
Customary diet and behavior counseling	NB32	622	109.3 (11.9)	-6.2 (0.3)	-3.4 (-4.3, -2.5)	<0.001
Study NB-304 (week 56)	Placebo	124	114.3 (12.4)	-2.9 (0.6)		
Obese subjects with type 2 diabetes	NB32	208	115.6 (12.6)	-5.0 (0.5)	-2.1 (-3.6, -0.6)	0.006

Sources:
Study NB-301: Clinical Study Report, Table 16
Study NB-302: Clinical Study Report, Table 14
Study NB-303: Clinical Study Report, Table 18
Study NB-304: Clinical Study Report, Table 14.2-52

TABLE 23 Phase 3 studies: Fasting HDL cholesterol (mg/dL), change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF)

	Group	n	Baseline mean (SD)	Percent Change from Baseline LSMean (SE)	LS Mean Difference from Placebo (95% CI)	p
Study NB-301 (week 56)	Placebo	345	52.0 (13.6)	-0.1 (0.5)		
Customary diet and	NB16	333	52.3 (13.4)	3.4 (0.5)	3.4 (2.2, 4.7)	<0.001
behavior counseling	NB32	359	51.9 (13.6)	3.4 (0.5)	3.5 (2.3, 4.7)	<0.001
Study NB-302 (week 56)	Placebo	144	55.3 (12.9)	0.9 (0.8)		
Intensive lifestyle modification counseling	NB32	392	53.6 (13.5)	4.1 (0.5)	3.2 (1.5, 5.0)	<0.001
Study NB-303 (week 28)	Placebo	308	51.4 (13.1)	-1.4 (0.4)		
Customary diet and behavior counseling	NB32	625	51.4 (13.3)	1.2 (0.3)	2.6 (1.6, 3.6)	<0.001
Study NB-304 (week 56)	Placebo	135	46.1 (11.5)	-0.3 (0.6)		
Obese subjects with type 2 diabetes	NB32	222	46.2 (10.2)	3.0 (0.5)	3.3 (1.8, 4.8)	<0.001

Sources:
Study NB-301: Clinical Study Report, Table 17
Study NB-302: Clinical Study Report, Table 17
Study NB-303: Clinical Study Report, Table 19
Study NB-304: Clinical Study Report, Table 18

TABLE 24 Phase 3 studies: Fasting triglycerides, % change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF)

	Group	n	Baseline geometric mean (mg/dL)	Percent Change from Baseline LSMean	p
Study NB-301 (week 56)	Placebo	345	113.2	-3.1	
Customary diet and	NB16	333	118.1	-8.0	0.046
behavior counseling	NB32	359	116.0	-12.7	<0.001
Study NB-302 (week 56)	Placebo	144	104.6	-8.5	
Intensive lifestyle modification counseling	NB32	392	111.6	-16.6	0.004
Study NB-303 (week 28)	Placebo	308	113.4	-1.4	
Customary diet and behavior counseling	NB32	625	119.0	-6.3	0.007
Study NB-304 (week 56)	Placebo	135	165.6	-0.8	
Obese subjects with type 2 diabetes	NB32	222	143.3	-11.2	0.007

Note:
The analysis was conducted on the logarithm transform of triglycerides. The geometric mean is the back-transformed mean on the logarithm scale. The change from baseline on the logarithm scale is back-transformed to a percentage change from baseline.

Sources:
Study NB-301: Clinical Study Report, Table 18
Study NB-302: Clinical Study Report, Table 15
Study NB-303: Clinical Study Report, Table 20
Study NB-304: Clinical Study Report, Table 17

TABLE 25 Phase 3 studies: Fasting triglycerides (mg/dL), change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF)

	Group	n	Baseline mean (SD)	Change from Baseline LSMean (SE)	LS Mean Difference from Placebo (95% CI)	p
Study NB-301 (week 56)	Placebo	345	127.4 (67.1)	-3.5 (3.1)		
Customary diet and	NB16	333	130.8 (62.3)	-9.3 (3.1)	-5.8 (-14.1, 2.5)	0.173
behavior counseling	NB32	359	129.6 (66.8)	-18.1 (3.0)	-14.6 (-22.8, -6.4)	<0.001
Study NB-302 (week 56)	Placebo	144	115.6 (56.4)	-11.3 (3.8)		
Intensive lifestyle modification counseling	NB32	392	126.2 (71.4)	-22.2 (2.4)	-10.8 (-18.9, -2.8)	0.009
Study NB-303 (week 28)	Placebo	308	128.7 (71.1)	-4.3 (2.9)		
Customary diet and behavior counseling	NB32	625	131.6 (61.5)	-8.5 (2.1)	-4.1 (-10.9, 2.6)	0.231
Study NB-304 (week 56)	Placebo	135	184.3 (89.3)	7.2 (8.3)		
Obese subjects with type 2 diabetes	NB32	222	165.0 (113.4)	-16.5 (6.5)	-23.8 (-42.6, -4.9)	0.014

Note: The analysis was conducted on the original scale of measurement of triglycerides (mg/dL).

Sources: Analysis by this reviewer

TABLE 26 Phase 3 studies: Impact of weight on quality of life (IWQOL-Lite) total score, change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF)

	Group	n	Baseline mean (SD)	Percent Change from Baseline LSMean (SE) (see notes)	LS Mean Difference from Placebo (95% CI) (see notes)	p
Study NB-301 (week 56)	Placebo	468	71.8 (17.2)	8.6 (0.5)		
Customary diet and	NB16	422	70.7 (17.0)	11.7 (0.5)	3.1 (1.7, 4.5)	<0.001
behavior counseling	NB32	417	70.3 (16.5)	12.7 (0.5)	4.1 (2.7, 5.6)	<0.001
Study NB-302 (week 56)	Placebo	178	101.9 (15.0)	-12.8 (1.1)		
Intensive lifestyle modification counseling	NB32	448	100.7 (15.4)	-16.7 (0.7)	-3.9 (-6.3, -1.5)	0.001
Study NB-303 (week 28)	Placebo	317	72.9 (15.7)	6.2 (0.6)		
Customary diet and behavior counseling	NB32	628	72.0 (17.4)	9.9 (0.4)	3.8 (2.5, 5.1)	<0.001
Study NB-304 (week 56)	Placebo	153	73.5 (16.9)	7.9 (0.9)		
Obese subjects with type 2 diabetes	NB32	241	73.2 (17.2)	9.3 (0.7)	1.4 (-0.8, 3.5)	0.208

Notes:

- For Studies NB-301, NB-303 and NB-304: The total score was transformed into 0 (worse) to 100 (best) using the formulae provided in the statistical analysis plans. An increase in the score correlates to an improvement in the quality of life.
- For Study NB-302: The total score was presented as raw, untransformed data and a decrease in the score correlates to an improvement in quality of life.

Sources:
Study NB-301: Clinical Study Report, Table 19
Study NB-302: Clinical Study Report, Table 18
Study NB-303: Clinical Study Report, Table 21
Study NB-304: Clinical Study Report, Table 14.2-54

TABLE 27 Study NB-304 (diabetic subjects): HbA1c, change from baseline to week 56 endpoint Full Analysis Set (LOCF)

Group	n	Baseline mean (SD)	Change from Baseline LSMean (SE)	LS Mean Difference from Placebo (95% CI)	p
Placebo	137	8.0 (0.9)	-0.1 (0.1)		
NB32	222	8.0 (0.8)	-0.6 (0.1)	-0.5 (-0.7, -0.3)	<0.001

Note: The analysis was conducted on the original scale of measurement of triglycerides (mg/dL).

Sources: Clinical Study Report, Table 14.2-34

3.2 Evaluation of Safety

A full evaluation of the safety of Contrave is included in the clinical review by Dr. Eileen Craig. Dr. Xiao Ding, Ph.D., a statistical reviewer from DB7, evaluated specific safety issues, including changes in blood pressure and heart rate. Contrave was the topic of a meeting of the Endocrine and Metabolic Drug Advisory Committee on December 7, 2010. A key topic of the meeting was the balance of risk and benefit to the target population of obese adults for whom Contrave may be prescribed. Central to the balance of risk and benefit was the effect of Contrave on blood pressure and heart rate. In this section, I provide a brief summary of results from the analysis of blood pressure, to illustrate some of these issues of concern.

The effect of bupropion in increasing systolic and diastolic blood pressure and heart rate has been well characterized and is described in the bupropion prescribing information. As was apparent in the analysis of secondary efficacy endpoints, the results for systolic blood pressure and diastolic blood pressure did not support the efficacy of Contrave, because the placebo-adjusted effect of Contrave was in the direction of an increase in mean SBP and DBP (TABLE 28, TABLE 29). The difference between the Contrave and placebo groups appeared by study week 4 and was maintained throughout the 56-week treatment period (FIGURE 10). The average decrease in blood pressure in the placebo group in each study reflects, in part, the improvement in blood pressure that is typically associated with weight loss. Based on the action of bupropion, it appears that Contrave counteracts this improvement in blood pressure to an extent that is related to weight loss. This overall dynamic creates a question about the overall balance of risk and benefit of Contrave, because an improvement in blood pressure is typically associated with a reduction in cardiovascular mortality, which is a long-term outcome of reducing obesity. This topic is addressed in more detail by Dr. Craig and Dr Xiao in their reviews. I discuss the

statistical issues involved in the evaluation of blood pressure changes in subgroups defined by weight loss in Part 4.3 of this review.

TABLE 28 Phase 3 studies: Systolic blood pressure (mm Hg), change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF)

	Group	n	Baseline mean (SD)	Percent Change from Baseline LSMean (SE)	LS Mean Difference from Placebo (95% CI)	p
Study NB-301 (week 56)	Placebo	511	119.0 (9.8)	-1.9 (0.4)		
Customary diet and	NB16	471	119.5 (9.9)	0.3 (0.4)	2.2 (1.2, 3.3)	<0.001
behavior counseling	NB32	471	118.9 (9.8)	-0.1 (0.4)	1.8 (0.8, 2.9)	<0.001
Study NB-302 (week 56)	Placebo	193	116.7 (10.9)	-3.9 (0.7)		
Intensive lifestyle modification counseling	NB32	482	116.9 (9.9)	-1.3 (0.5)	2.6 (1.0, 4.1)	0.002
Study NB-303 (week 28)	Placebo	456	118.2 (10.5)	-1.2 (0.4)		
Customary diet and behavior counseling	NB32	824	118.1 (10.0)	-0.9 (0.3)	0.3 (-0.7, 1.3)	0.556
Study NB-304 (week 56)	Placebo	159	124.5 (9.6)	-1.1 (0.9)		
Obese subjects with type 2 diabetes	NB32	265	125.0 (11.0)	0.0 (0.7)	1.1 (-1.0, 3.3)	0.297

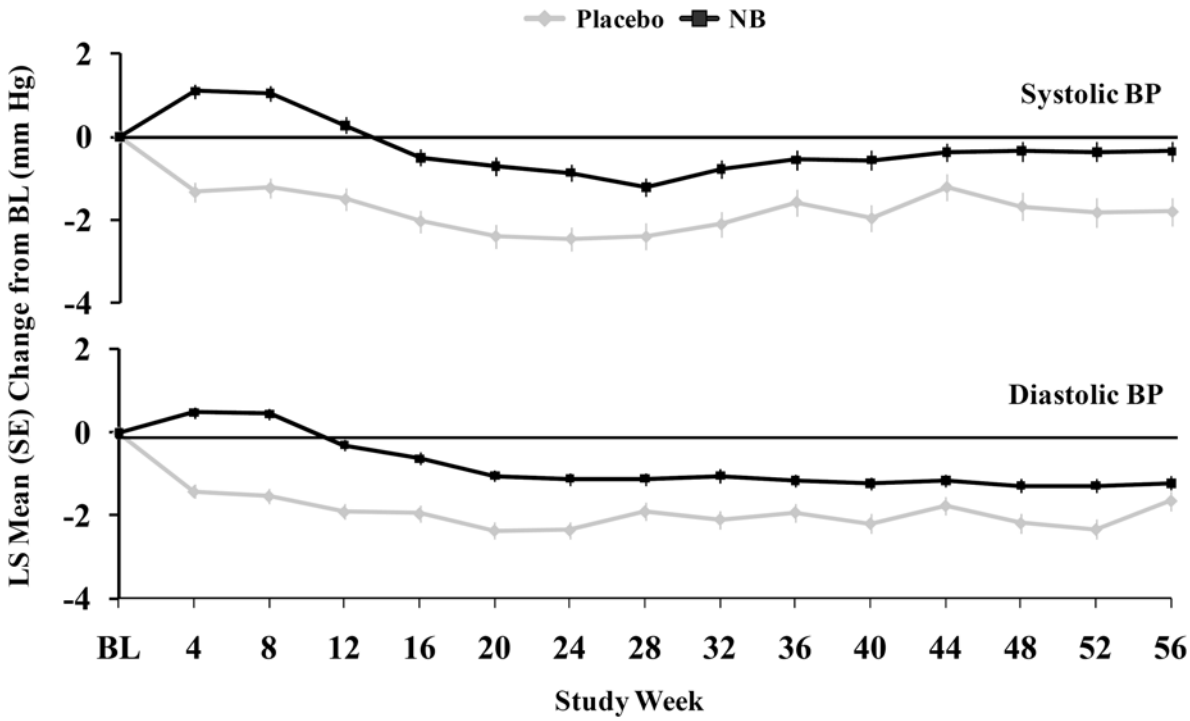
Sources:
Study NB-301: Clinical Study Report, Table 14.2-50
Study NB-302: Clinical Study Report, Table 14.2-56
Study NB-303: Clinical Study Report, Table 14.2-91
Study NB-304: Clinical Study Report, Table 14.2-59

TABLE 29 Phase 3 studies: Diastolic blood pressure (mm Hg), change from baseline to week 56 endpoint (week 28 for Study NB-303); Full Analysis Set (LOCF)

	Group	n	Baseline mean (SD)	Percent Change from Baseline LSMean (SE)	LS Mean Difference from Placebo (95% CI)	p
Study NB-301 (week 56)	Placebo	511	77.3 (6.6)	-0.9 (0.3)		
Customary diet and	NB16	471	76.6 (7.2)	0.1 (0.3)	1.0 (0.2, 1.7)	0.015
behavior counseling	NB32	471	77.1 (7.2)	0.0 (0.3)	0.9 (0.1, 1.7)	0.022
Study NB-302 (week 56)	Placebo	193	77.2 (7.4)	-2.8 (0.5)		
Intensive lifestyle modification counseling	NB32	482	78.2 (7.2)	-1.4 (0.3)	1.4 (0.3, 2.5)	0.017
Study NB-303 (week 28)	Placebo	456	76.8 (7.0)	-0.7 (0.3)		
Customary diet and behavior counseling	NB32	824	76.8 (7.0)	0.2 (0.2)	0.9 (0.2, 1.6)	0.017
Study NB-304 (week 56)	Placebo	159	77.4 (7.1)	-1.5 (0.6)		
Obese subjects with type 2 diabetes	NB32	265	77.5 (7.5)	-1.1 (0.5)	0.4 (-1.0, 1.9)	0.582

Sources:
Study NB-301: Clinical Study Report, Table 4.2-52
Study NB-302: Clinical Study Report, Table 14.2-59
Study NB-303: Clinical Study Report, Table 14.2-83
Study NB-304: Clinical Study Report, Table 14.2-61

FIGURE 10 Blood pressure (mm Hg), repeated measures analysis of change from baseline to each visit: Primary safety dataset, double blind treatment phase



Source: Orexigen, briefing materials for October 13, 2010 meeting, Figure 4

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Sex, Race, Ethnicity and Age

Sex: Females made up the large majority of each study (about 75%). However, the studies were large enough to evaluate the placebo-adjusted effect of lorcaserin in males and females. In all four Phase 3 studies, males and females were fairly similar in the mean placebo-adjusted effect of NB32 in the proportion of subjects who were 5% responders (FIGURE 11A). The interaction between treatment arm and sex was not statistically significant in the logistic regression analysis from any of the Phase 3 studies. The percentage of male subjects who completed a study was somewhat lower than the percentage of female subjects, by study and study arm.

Race and Ethnicity: Subjects in the Caucasian/White race subgroup and in the non-Hispanic/Latino ethnicity subgroup made up the large majority of each study (about 75% in the Caucasian/White subgroup and about 88% in the non-Hispanic/Latino subgroup). However, the studies were large enough to describe the placebo-adjusted effect of lorcaserin in African

American/Black, and Hispanic/Latino subgroups. Subjects in the Caucasian/White subgroup were fairly similar to subjects in the African American/Black subgroup, and non-Hispanic/Latino subjects were fairly similar to Hispanic/Latino subjects, with respect to the placebo-adjusted effect of NB32 (FIGURE 11B and C). However, the unadjusted mean weight loss in the placebo and NB32 arms was less in the African American/Black subgroup compared to the Caucasian/White subgroup, in Studies NB-301, NB-302 and NB-303. This finding corresponds to a greater percentage of subjects in the African American/Black subgroup who discontinued study medication compared to subjects in the Caucasian/White subgroup in each arm of these three studies. This pattern was not apparent in the study that enrolled diabetics (Study NB-304), where the unadjusted mean weight loss in the placebo and the NB32 arms, and the percentage of subjects who discontinued study medication was fairly similar between the African American/Black subgroup and the Caucasian/White subgroup.

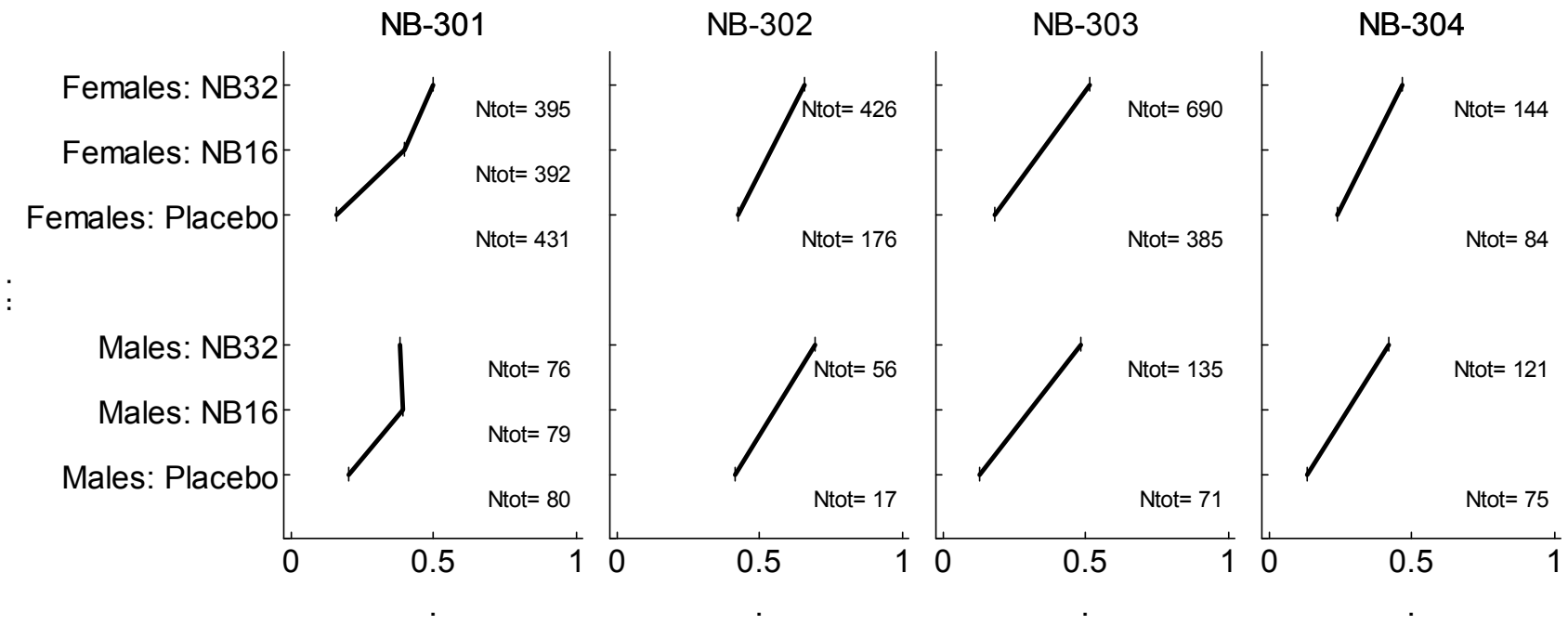
Age: The enrollment criteria in the three studies that enrolled non-diabetic subjects (Study NB-301, Study NB-302 and Study NB-303) excluded subjects who were over 65 years old. A small percentage of subjects (12%) in the study that enrolled diabetic subjects (Study NB-304) were 65 and over. For this reason, I did not evaluate the comparative effect of NB32 in this older ae group.

4.2 Other Special/Subgroup Populations Defined by Medical Conditions at Baseline

The placebo-adjusted effect of NB32 on the proportion of 5% responders was relatively similar among subgroups defined either by BMI at baseline, or by the presence or absence of hypertension, dyslipidemia, or metabolic syndrome at baseline in each study (FIGURE 12). A statistically significant interaction between treatment group and subgroup defined by baseline metabolic syndrome was apparent in study NB-301 ($p=0.080$; where $p=0.1$ is used to screen for statistical significance among treatment group by subgroup interaction terms). However, this relationship was not apparent in the other three studies, and it did not appear to stem from a lack of parallelism in the dose response of each subgroup ((FIGURE 12D).

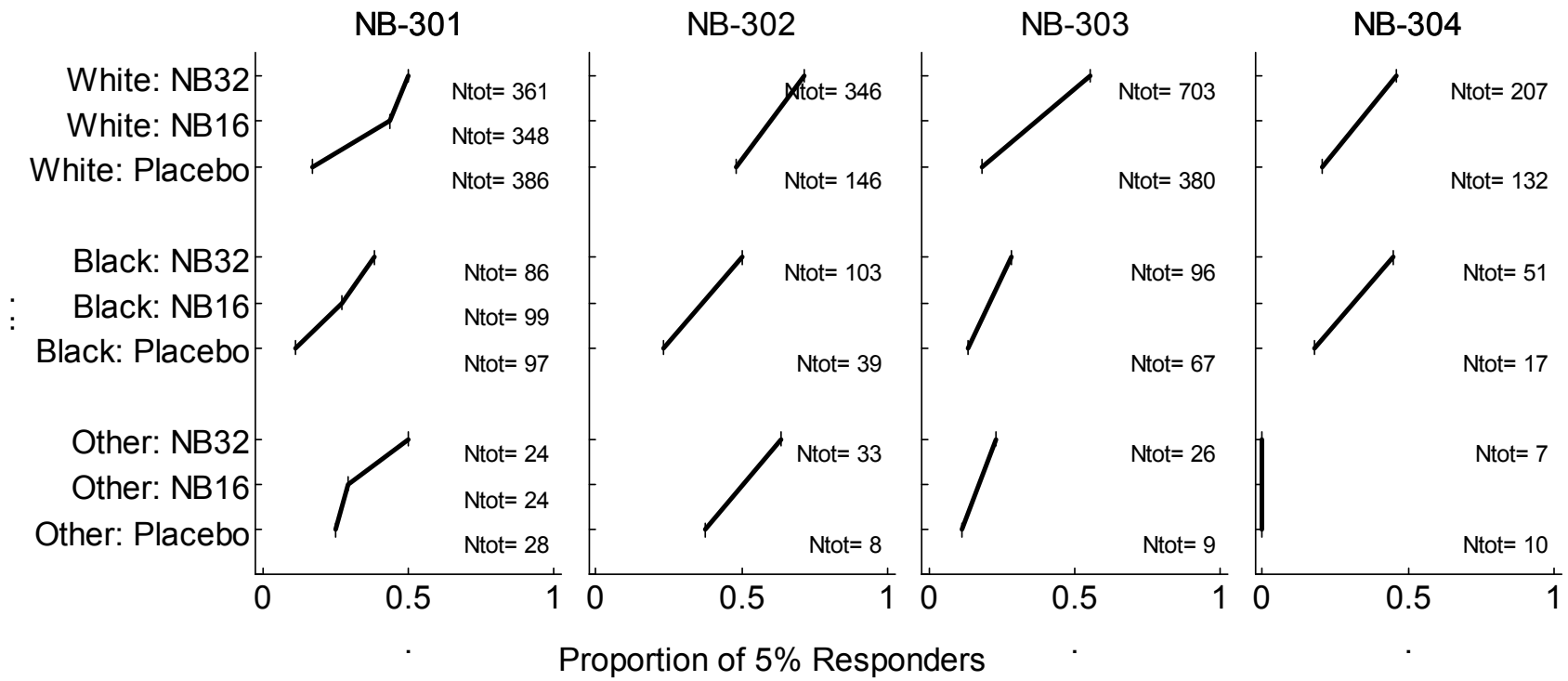
FIGURE 11 The proportion of 5% responders in subgroups (FAS/LOCF): Gender, race and ethnicity

A. Gender



sex x treatment group:	p=0.160		p=0.728		p=0.406		p=0.276	
Table reference	Table 14.2-22		Table 14.2-21		Table 14.2-45		Table 14.2-24	
Percentage who discontinued study med.	Females	Males	Females	Males	Females	Males	Females	Males
NB32	39.2%	26.3%	60.3%	27.3%	35.1%	33.3%	38.2%	28.9%
NB16	41.6%	30.4%						
Placebo	44.3%	60.0%	44.6%	89.9%	43.6%	29.6%	41.7%	32.0%

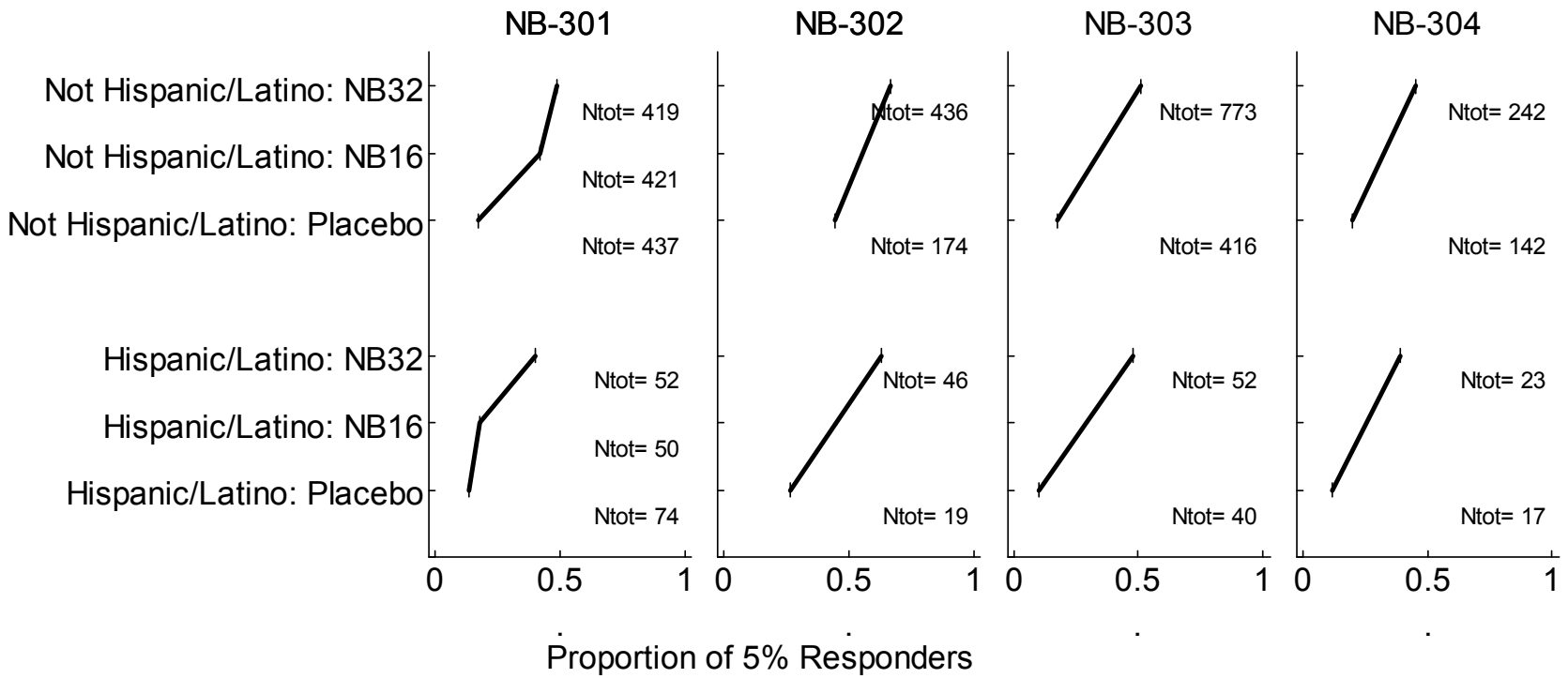
B. Race (White, Black, Other races combined)



race by group:	p=0.333	p=0.614	p=0.246	p=0.472
Table reference:	Table 14.2-23	Table 14.2-22	Table 14.2-46	Table 14.2-25

Percentage who discontinued study med.	White	Black	White	Black	White	Black	White	Black
NB32	36.3%	40.7%	33.5%	50.5%	32.4%	45.8%	30.9%	43.1%
NB16	39.4%	40.4%						
Placebo	43.0%	43.3%	43.2%	51.3%	40.3%	46.3%	37.1%	35.3%

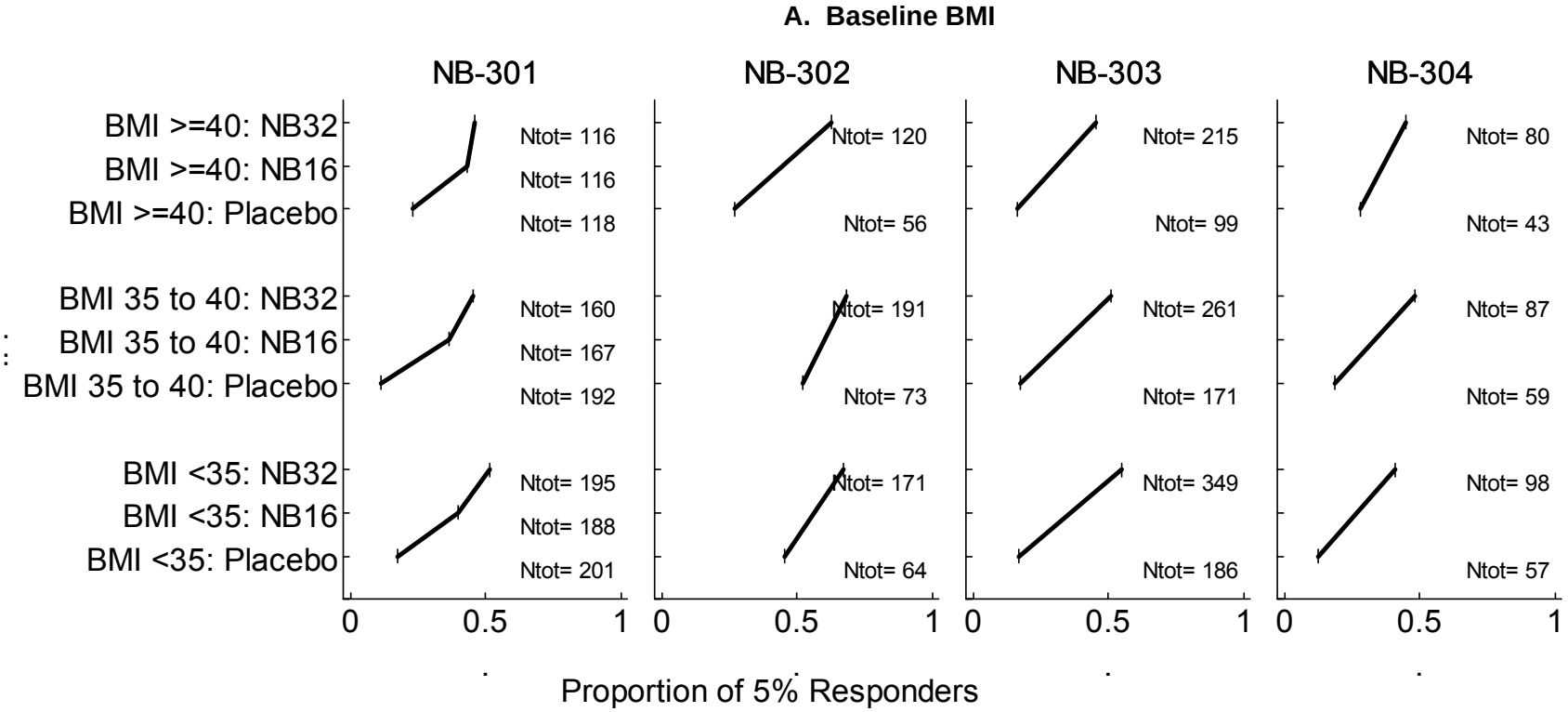
C. Ethnicity (Hispanic/Latino, not Hispanic Latino)



	NB-301		NB-302		NB-303		NB-304	
ethnicity by treatment group: (analysis by this reviewer)	p=0.133		p=0.326		p=0.378		p=0.649	
	Not H/L	H/L	Not H/L	H/L	Not H/L	H/L	Not H/L	H/L
NB32	34.8%	55.8%	37.4%	39.1%	34.7%	36.5%	33.9%	34.8%
NB16	38.5%	50.0%						
Placebo	40.5%	59.5%	43.7%	57.9%	40.1%	55.0%	35.9%	47.1%

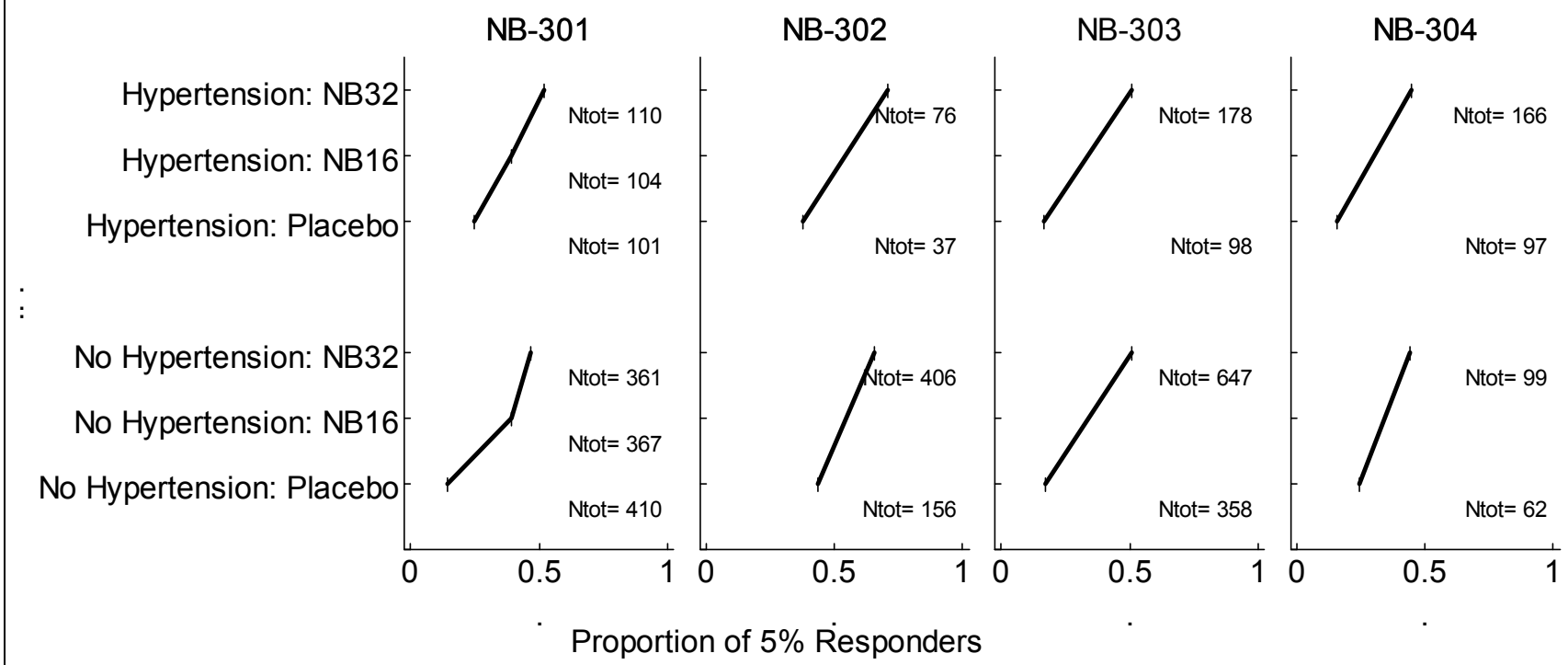
Note: The proportion of 5% responders for each subgroup category and treatment group is depicted on the graph. The p-values are from the logistic regression analysis with the following general form: treatment group, subgroup, baseline body weight, treatment by subgroup interaction. The race had two subgroups; white and other.

FIGURE 12 The proportion of 5% responders in subgroups defined by medical condition at baseline (FAS/LOCF)



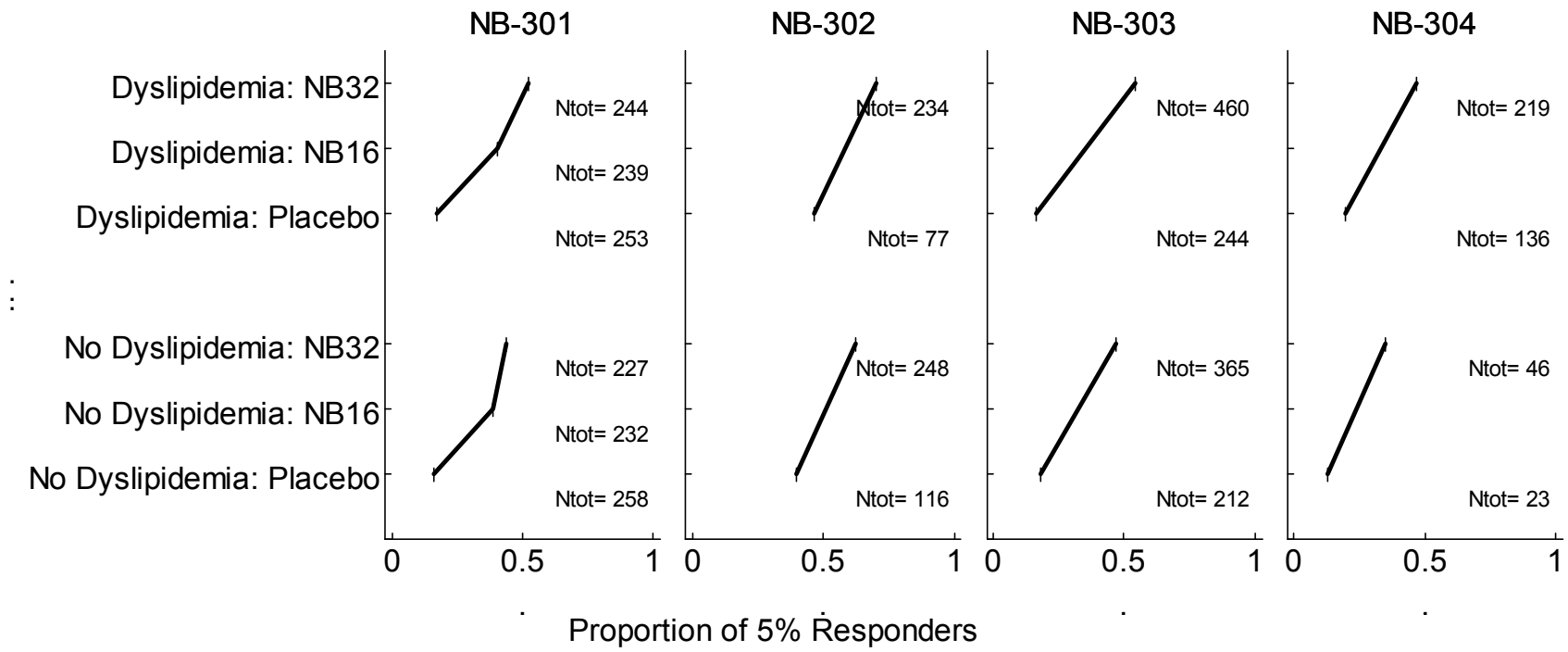
BMI category x treatment group:	p=0.370	p=0.503	p=0.716	p=0.369
Table reference:	Table 14.2-26	Table 14.2-24	Table 14.2-49	Table 14.2-28

B. Hypertension at Baseline



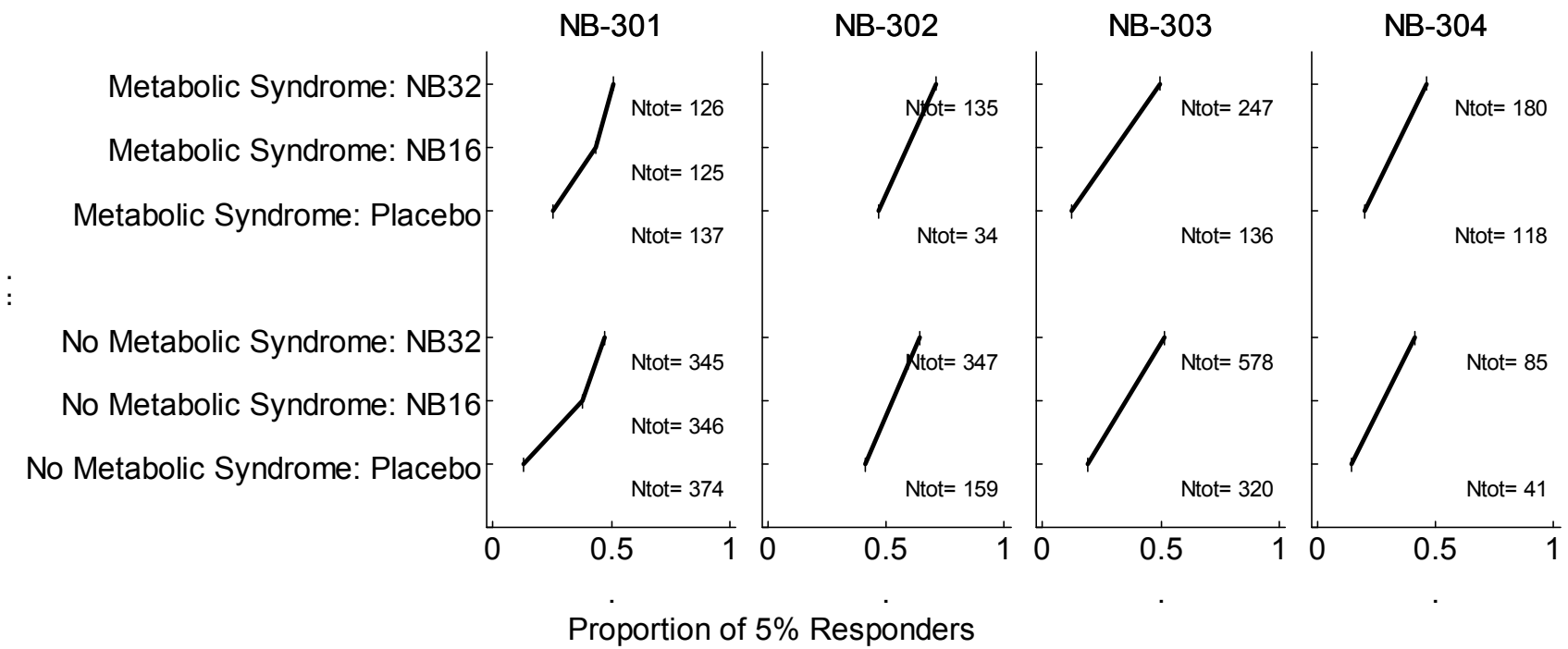
hypertension by treatment group:	p=0.165	p=0.254	p=0.907	p=0.232
Table reference:	Table 14.2-27	Table 14.2-25	Table 14.2-50	Table 14.2-29

C. Dyslipidemia at Baseline



dyslipidemia by treatment group:	p=0.498	p=0.826	p=0.544	p=0.938
Table reference:	Table 14.2-28	Table 14.2-26	Table 14.2-51	Table 14.2-30

D. Metabolic Syndrome at Baseline



metabolic syndrome by treatment group:	NB-301	NB-302	NB-303	NB-304
Analysis by this reviewer	p=0.080	p=0.863	p=0.211	p=0.743

Notes: The proportion of 5% responders for each subgroup category and treatment group is depicted on the graph. The p-values are from the logistic regression analysis with the following general form: treatment group, subgroup, baseline body weight, treatment by subgroup interaction. The BMI subgroup had two categories, separated at the median BMI of the study population.

4.3 Subgroups Defined by Responder Status

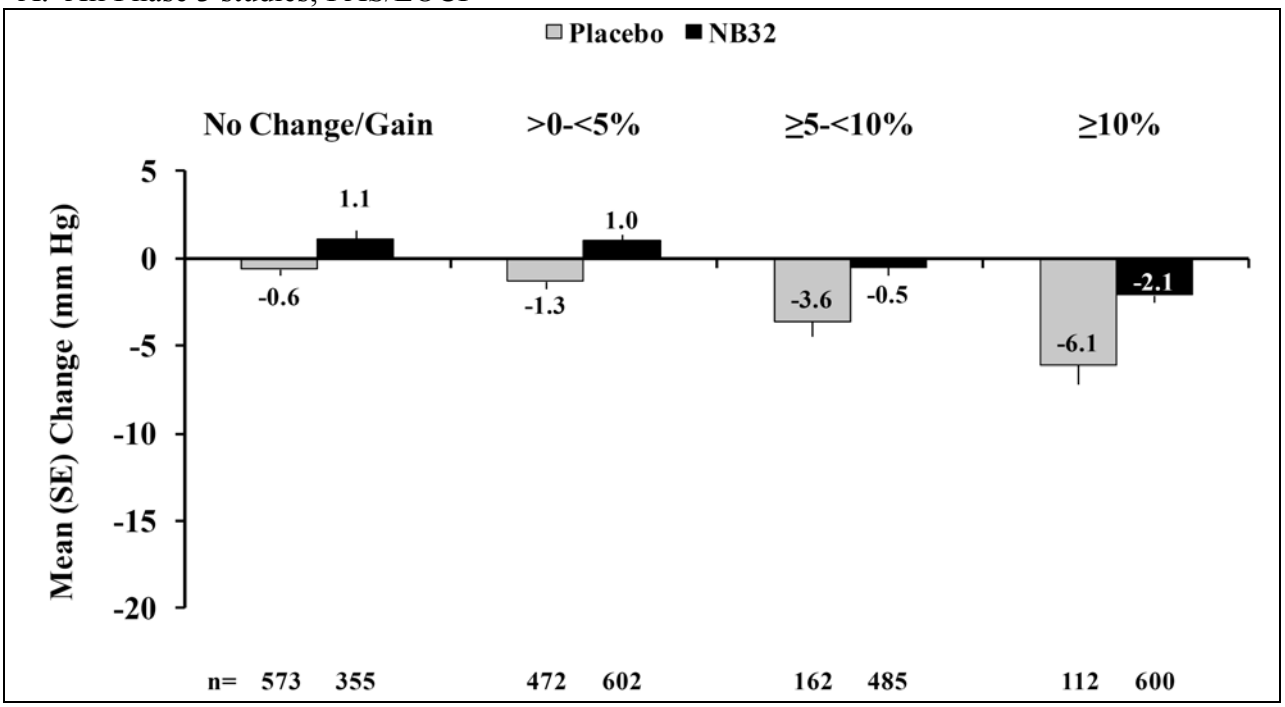
4.3.1 Balance of risk and benefit: Weight loss and blood pressure change in Contrave and placebo groups

Because one of the goals of weight loss is to reduce the risk of cardiovascular mortality, for which a reduction in blood pressure is an important biomarker, the Division requested additional exploratory evaluations of the relationship between weight loss and change in blood pressure in the placebo and Contrave groups. Among the requests were an evaluation of change in blood pressure in subgroups defined by treatment group and status with respect to weight loss at study endpoint. This was the topic of a meeting held on October 13, 2010 between Orexigen and the Agency.

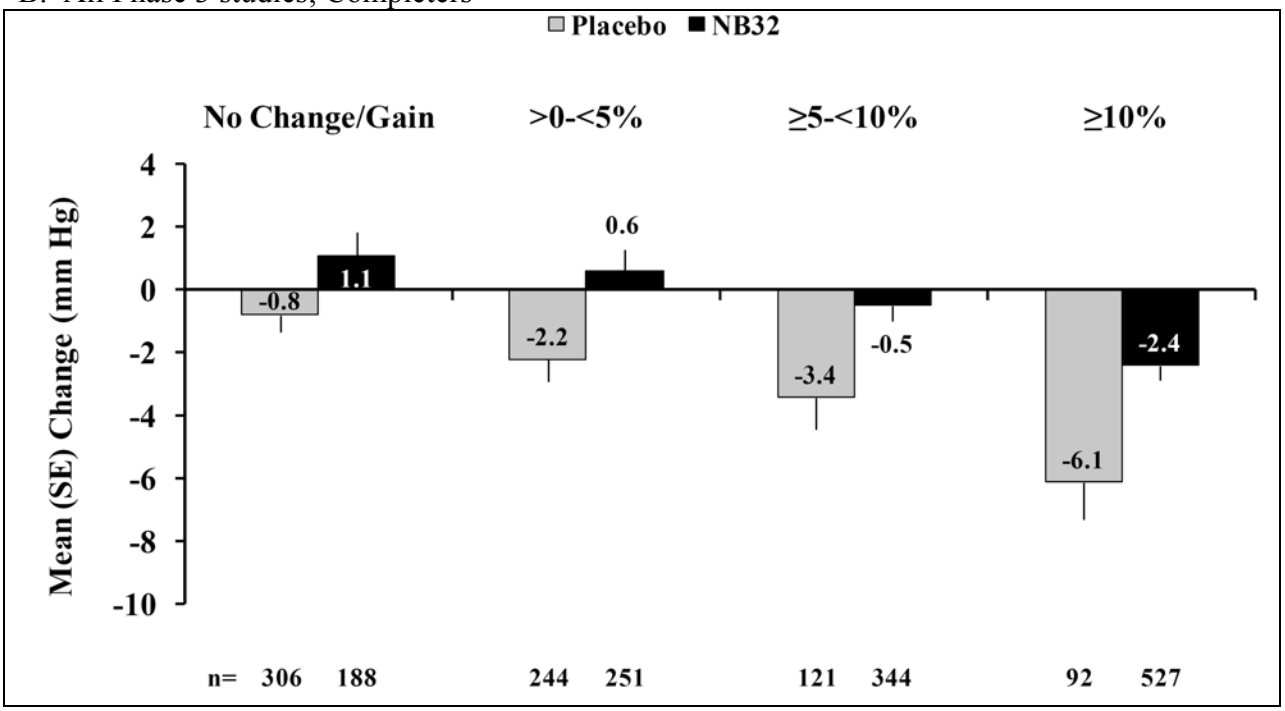
The summary findings are consistent with the interpretation that, to some extent, Contrave counteracts the beneficial relationship between weight loss and an improvement in blood pressure (FIGURE 13 - FIGURE 15). However, these findings are exploratory. A statistical comparison between NB32 and placebo within a subgroup, for example, subjects who had lost between 5% and 10% of baseline body weight, would be difficult to interpret in the context of a randomized clinical study because the subgroups are not comparable. However, given what is known about the blood pressure effect of bupropion, and the beneficial effect of losing weight on blood pressure, the patterns observed between weight loss and blood pressure change in the Contrave and placebo groups are credible.

FIGURE 13 Systolic blood pressure (mm Hg), change from baseline to week 56 endpoint for all Phase 3 studies (pooled) by weight loss category

A. All Phase 3 studies; FAS/LOCF

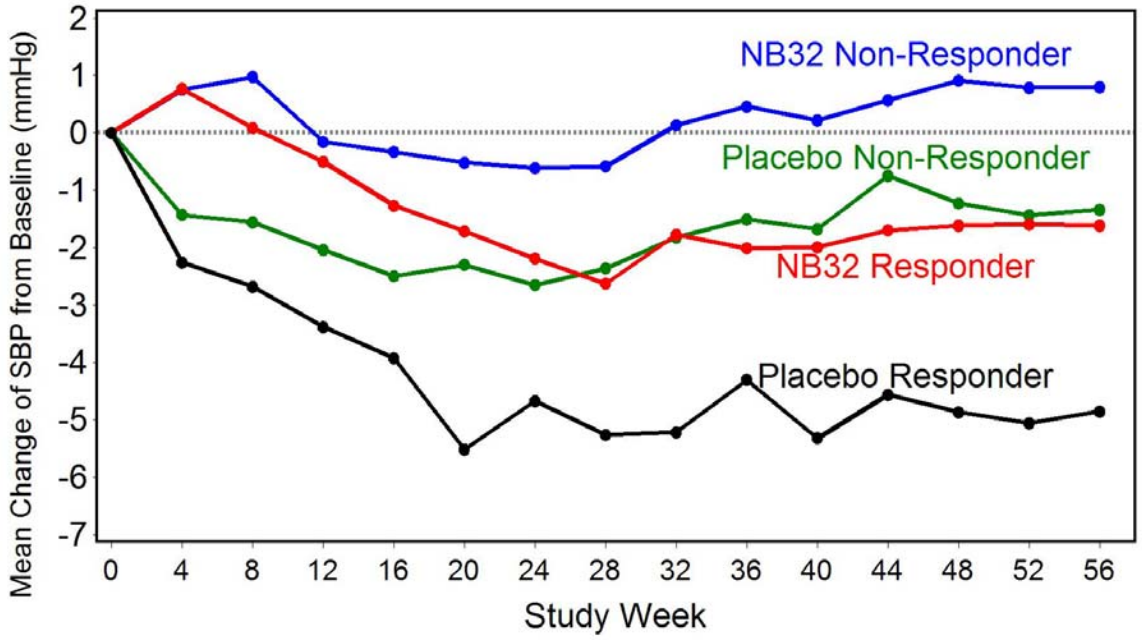


B. All Phase 3 studies; Completers



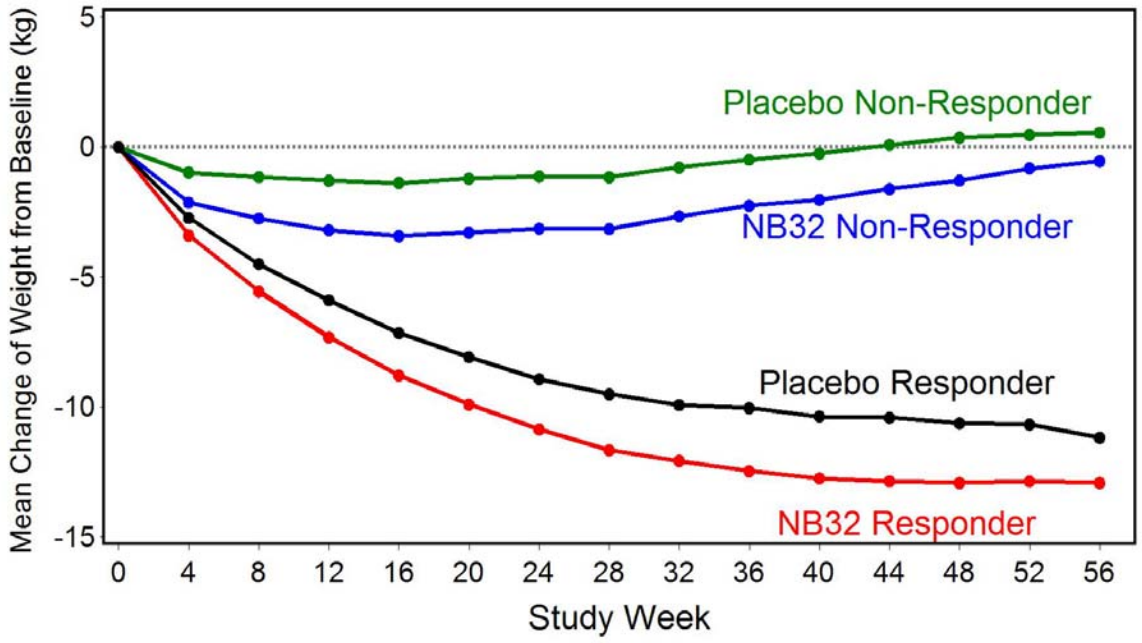
Source: Orexigen, briefing materials for October 13, 2010 meeting, Figure 8 and Figure 9

FIGURE 14 Systolic blood pressure (mm Hg), change from baseline to week 56 endpoint for all Phase 3 studies (pooled) by 5% responder category, completers population



Source: Analysis by Dr. Xiao Ding, DB7 statistical reviewer

FIGURE 15 Weight (kg), change from baseline to week 56 endpoint for all Phase 3 studies (pooled) by 5% responder category, completers population



Source: Analysis by Dr. Xiao Ding, DB7 statistical reviewer

4.3.2. Risk Management: Predicting week 56 weight loss and blood pressure change from early results

As part of the October 13, 2010 meeting with the Division, the applicant proposed a patient management algorithm for identifying patients who may not benefit from Contrave, either due to blood pressure elevation or insufficient weight loss or both (FIGURE 16). In support of this algorithm, the applicant cited results from an exploratory analysis of the Phase 3 studies that demonstrated that early weight loss is a reasonable predictor of one year weight loss and that early blood pressure changes are a reasonable predictor of long-term blood pressure elevations (TABLE 30). I was not able to locate the analytical results that supported these findings in the submissions to the NDA. However, if the Division views this type of approach as useful, I have the following recommendations:

1. The Division can review the criteria for the clinical importance of blood pressure changes and weight changes and the time frame for evaluation, and propose revisions, if needed.

2. Based on the proposed revisions, the applicant can do the following:
    a. Obtain and evaluate prediction equations for each Phase 3 study separately.
    b. For each study, use the intention-to-treat population, with non-responder imputation for subjects who discontinued study medication prior to the week at which the predictive value of early weight loss or early blood pressure changes are to be evaluated.
    c. Assess whether results from each study are similar enough to be combined; interpret any substantial differences among studies.
    d. If study results can be combined, obtain combined prediction equations for weight loss and blood pressure changes. Develop a patient management algorithm based on the combined prediction equations.
    e. Submit the proposed patient management algorithm and the analytical results to support the algorithm for review by the Division and the biometrics team.

TABLE 30 Proposed Patient Management Algorithm from Orexigen

The Patient Management Algorithm (see [FIGURE 16]) which is intended to help inform patient selection and management decisions, has been developed based on analyses of data from the Contrave clinical development program. Based on these analyses, two key variables emerged as the most relevant to inform appropriate patient management, specifically (1) early weight loss as a predictor of one year weight loss and (2) early blood pressure changes as a predictor of long-term blood pressure elevations. As seen in [FIGURE 16] and presented in more detail below, it is these two variables that comprise the main elements of the algorithm and support the use of a therapeutic trial.

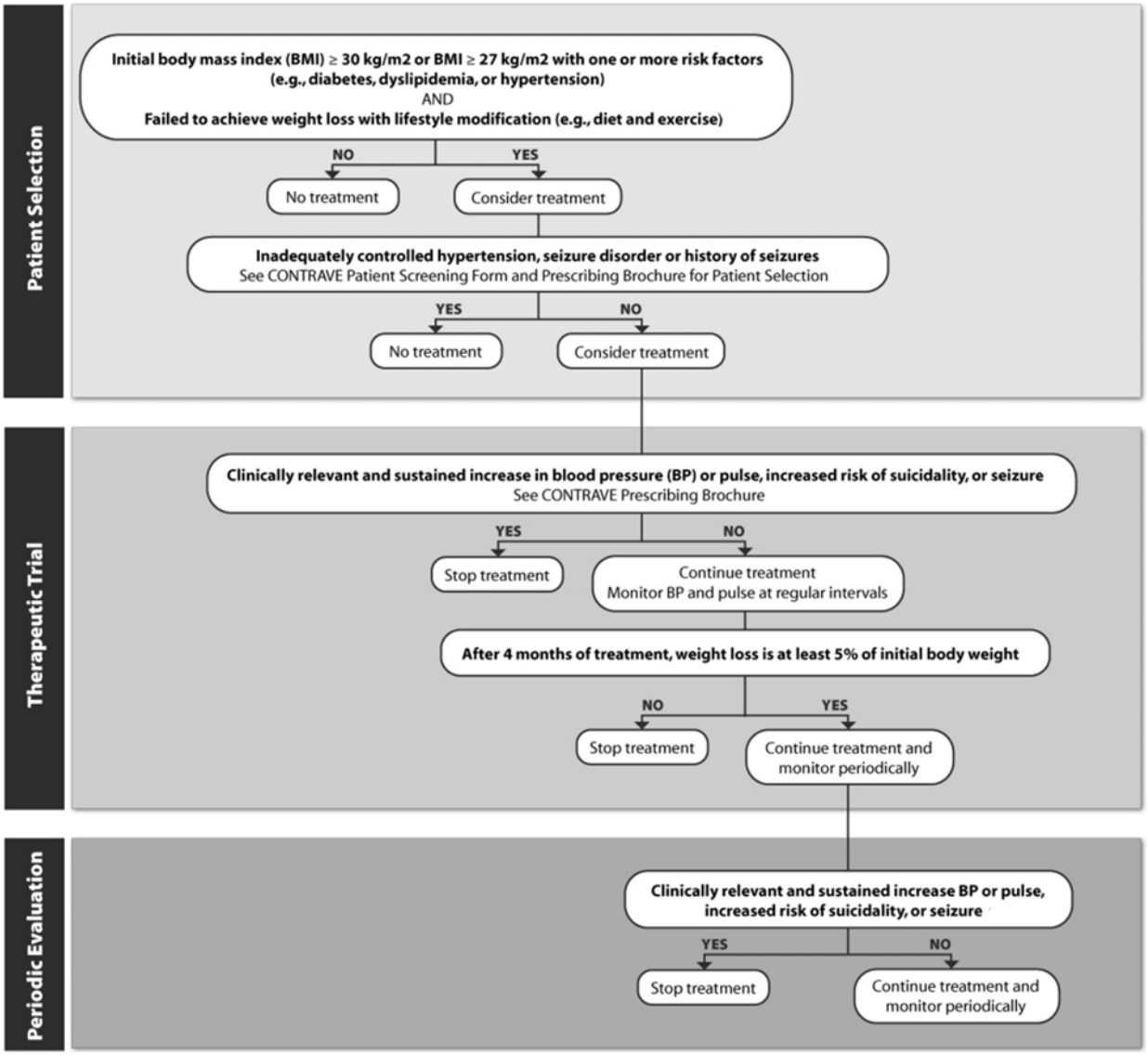
Weight Reduction. Analyses were conducted to evaluate whether earlier weight loss (Weeks 4-28) is predictive of a 5% or greater weight loss response at week 56. Based on receiver operating characteristic curves 5% weight loss from baseline at Week 16 showed 75 to 85% accuracy in the four Phase 3 trials in identifying 5% responders at Week 56 with fair balance between sensitivity and specificity. Additionally, in a pooled analysis of the four Phase 3 trials, among the NB32 subjects who achieved ≥ 5% weight loss at Week 56 based on LOCF, more than 85% reached the responder status by Week 16.

Blood Pressure and Heart Rate. Odds ratios based on logistic modeling were used to assess whether the occurrence of late outliers (Weeks 28-56) of blood pressure and heart rate could be predicted based on the occurrence of early outliers (Weeks 4-16). For this assessment outliers were defined as at least 2 consecutive increases in the vital sign parameter ≥ 10 units relative to baseline during the early and/or late time period. Patients with an earlier SBP outlier had at least 12 times the odds of having outlier values late in treatment, compared to those without early outliers. Similar findings were noted for DBP and heart rate.

Based on these analyses it is anticipated that close monitoring and management of these parameters may be useful in selecting appropriate patients for continued Contrave treatment and in stopping therapy for patients developing meaningful blood pressure or HR increases, or achieving insufficient weight loss (<5%).

Source: Orexigen, briefing materials for October 13, 2010 meeting, Part 9.2.3.1
(the figure reference was changed to refer to the figure in this review)

FIGURE 16 Proposed patient management algorithm



Source: Orexigen, briefing document for December 7, 2010 EMDAC meeting, Figure 36

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The applicant used the results from a Phase 2 Study NB-201 to support the contribution of naltrexone and bupropion separately to the overall efficacy of Contrave. A post-hoc Bonferroni correction was used to conclude that each component contributed to the efficacy of the combination. While the p-value was fairly low and did not cause great concern, this experience supported the need for a more careful statistical review of the pivotal study(ies) to be used to support the combination product.

A substantial percentage of randomized subjects in each study and study arm, between 41% and 51%, discontinued from taking study medication prior to week 56. This is typical of weight loss studies. This experience underscores the need to emphasize the importance of tracking all randomized subjects in a weight loss study, even those who discontinue study medication and/or stop attending the clinic. However, I believe it is reasonable to assume that subjects who dropped out this early were not likely to reach a 5% weight loss goal. Classifying subjects who discontinued early as non-responders on the 5% categorical endpoint permits the use of the intention-to-treat data base. In my opinion, this approach is most likely to represent the experience of the intended target population of Contrave.

5.2 Conclusions

The results of four Phase 3 studies are consistent and confirm the efficacy of naltrexone 32 mg /bupropion 360 mg (NB32) compared to placebo after 56 weeks of treatment in three studies and 28 weeks of treatment in one study. Results of alternate analysis models and other versions of the analysis population were consistent with the results from the primary analysis. The results from the analysis of key secondary efficacy endpoints, such as triglycerides, HDL-cholesterol, waist circumference, and the total score on the Impact of Weight on the Quality of Life questionnaire, support the efficacy of the NB32 compared to placebo.

The placebo-adjusted effect of NB32 on the proportion of 5% responders was fairly similar among subgroups defined by BMI at baseline, or by the presence or absence of hypertension, dyslipidemia, or metabolic syndrome at baseline in each study. Although the majority of subjects were female, white/Caucasian and not Hispanic, the studies were large enough to assess the effect of gender, white/Caucasian compared to black/African American and Hispanic compared to not Hispanic on the efficacy of NB32. The placebo-adjusted effect of NB32 was fairly similar across these demographic factors in each of the Phase 3 studies.

In the study that enrolled subjects with Type 2 diabetes, the mean HbA1c at week 52 was in the direction of improved glycemic control in the NB32 group compared to the placebo group.

The applicant has proposed an algorithm for the early identification of subjects who may not benefit from Contrave because of insufficient weight loss and/or an unacceptable increase in

blood pressure. Recommendations for developing the statistical aspects of this algorithm and for submitting this information for further review are included in part 4.3.2 of this review.

5.3 Recommendations for Labeling

I have the following recommendations for part 14 (Clinical Studies) of the package insert, based on the version that the applicant submitted with the NDA/0 on 3/31/10:

1. The introductory description of the Phase 3 studies should include a description of the percentage of subjects who discontinued study medication, the timing of this discontinuation and the most common reasons for discontinuation.

2. The discussion of results from the analysis population of (b) (4) occurs throughout in the text and in the tables, and I believe that this discussion and these results should be omitted. For example, the sentence (b) (4) should be omitted from the introduction to part 14. I recommend that the results for (b) (4) be omitted from Tables 4 and 5.

3. The applicant has reported the results from the FAS/LOCF analysis population from each of the phase 3 studies, using the terms (b) (4) and "LOCF" to refer to these results. I believe this is appropriate because this was the pre-specified primary analysis of the co-primary endpoints. The term (b) (4) is an established term for (b) (4) and "LOCF" is an established term for last observation carried forward. However, these acronyms should probably be written out in full at the first occurrence in the Clinical Studies section.

4. Results are presented for week 56 of Study NB-303 (b) (4) in the label). Week 28 was the primary endpoint period for Study NB-303. However, the week 56 results for the co-primary endpoints were identified as key secondary endpoints, and these endpoints were included in the set of endpoints with statistically significant results. This may support including results about week 56 from Study NB-303.

5. Summaries of the study results in Table 4 should include information about the percentage of subjects who discontinued study medication in each treatment arm. A comment about the percentage of study dropouts might be useful in the other tables as well.

6. Throughout the text and in the tables, summary statistics should be rounded to the nearest 0.1 decimal.

7. Table 5 summarizes the percentages of patients losing $\geq 5\%$, $\geq 10\%$ (b) (4) of body weight from baseline in 56 weeks in each study. I believe the 5% responder and the 10% responder have sufficient statistical support to be included in this table. The 5% responder is a co-primary endpoint. The 10% responder is a key secondary endpoint that was included in the statistically significant sequence of endpoints in all but Study NB-

304. [redacted]

8. Figure 1 describes the longitudinal profile of weight change in Study NB-301 ("COR-I"). The plot also depicts, using separate symbols, the Least Squares Means for the 56-week primary endpoint, based on the [redacted] LOCF analysis population. I agree with depicting the week 56 LSMeans as a separate result on this plot. In addition, I believe that this separate result for week 56 can include symbols for statistical significance, because these are the results from the primary efficacy analysis. Otherwise, I recommend that Figure 1 should be modified as follows: (a) use the completers analysis population for the longitudinal profile; [redacted] because these are exploratory and not primary; (c) include a description of the percentage of subjects who discontinued from taking study medication.

9. In the section on "Effect on Cardiovascular and Metabolic Parameters":

   a. Study results are reported in a subgroup of patients diagnosed with [redacted]. This was not a primary or key secondary analysis. This may need to be omitted.

   b. Study results should be reported separately for Studies NB-301 ("COR-I") and [redacted] in Table 6. [redacted]

   c. For HDL, LDL and TG, it may be more useful to report the mean differences from baseline on the original scale (mg/dL) of measurement.

   d. Table 6 summarizes the statistical results from key secondary endpoints. The Division may be considering which endpoints to include in this table. I suggest that the p-values be reported with some explanation, perhaps in a footnote, of the protection of Type I error given to these key endpoints by pre-specifying the sequence of testing.

10. In the section on "Effects on [redacted]

    a. A statistically significant result is described for week 16. I recommend that this result be omitted because analyses at early time points were exploratory.

    b. A result is described for patients with baseline [redacted] I recommend that this result be omitted because analysis for this subgroup was exploratory.

    c. Table 7 reports results for key secondary endpoints related to glycemic control. The Division may be considering which endpoints to include in this table. I suggest that the p-values be reported with some explanation, perhaps in a footnote, of the protection of Type I error given to these key endpoints by pre-specifying the sequence of testing. For example, only HbA1c (change from

baseline) was included in the sequence of statistically significant secondary endpoints. Other endpoints, such as the percentage of patients achieving HbA1c < 7% and the percentage of patients requiring rescue diabetes medications, had nominal p-values < 0.05 but in fact were not included in the sequence of statistically significant secondary endpoints.

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

/s/

JANICE A DERR
12/15/2010

JON T SAHLROOT
12/15/2010
concur

THOMAS J PERMUTT
12/15/2010
concur

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 200063 **Applicant:** Orexigen **Received:** March 31, 2010

Drug Name: Contrave extended-release tablets (naltrexone / bupropion) **NDA/BLA Type:** standard **Filing meeting:** May 24, 2010

Date of Statistical filing checklist: 5/11/10

On **initial** overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	√			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	√			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).	√			
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).	√			Study data uses ADaM and STDM data standards. Define.xml files are provided, and annotation is comprehensive and understandable. Disposition variables are complete.

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? ___Yes_____

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter. See the following:

We request a summary of changes to specific tables and figures in each document that took place as a result of the corrections to the efficacy endpoints in the clinical study NB-303, as described in your submission 0001 dated May 4, 2010.

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

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

Other summary comments (These are for internal purposes only and are not to be transmitted to the sponsor).

Phase 3 studies:

Study NB-301: Placebo (n=581), Naltrexone 16mg/day + Bupropion 360mg/day (NB16; n=578), Naltrexone 32mg/day and Bupropion 360mg/day (NB32; n-583); 56-week double-blind; obese subjects 18-66 yrs old, with or without controlled hypertension and/or dyslipidemia

Study NB-302: Placebo (n=202) and NB32 (n=591); 56-week double-blind, obese subjects, 19-65 yrs old, with or without controlled hypertension and/or dyslipidemia; nonsmokers; participated in intense group lifestyle modification counseling (28 sessions)

Study NB-303: Placebo (n=495), NB32 (n=1001); 56-week double-blind*, obese subjects 18-65 yrs old with or without controlled hypertension and/or dyslipidemia; primary efficacy evaluation was conducted at week 28 with secondary evaluation at week 56. *At week 28 non-responders were unblinded and the NB32 non-responders had a dosage increase to NB48; this applied to 123 subjects (placebo group non-responders stayed on placebo).

Study NB-304: Placebo (n=170), NB32 (n=335); 56-week double-blind, obese subjects 20 to 72 yrs old with type 2 diabetes and with or without controlled hypertension and/or dyslipidemia.

Phase 2 combination study NB-201: placebo, monotherapy and combination therapy arms in a range of doses; 24-week double-blind followed by 24-week extension; obese subjects 18-60 yrs old without complicated obesity who are nonsmokers.

Submission 0001 rec'd 5/4/01 is a clinical amendment that affects NB-303, correcting a mistake in efficacy calculations that affect the week 56 endpoints (for Study NB-303, the corrected endpoints pertain to week 56 which is after the 28-week primary endpoint). The amendment includes a revised study report and a revised ISE. Orexigen offers to provide a full summary of the changes this correction made to tables and figures in the submission, and I recommend that we do ask for this.

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-200063	ORIG-1	OREXIGEN THERAPEUTICS INC	CONTRAVE® (Naltrexone HCl and Bupropion HCl)

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

/s/

JANICE A DERR
05/10/2010

JON T SAHLROOT
05/11/2010