CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

021876Orig1s000

STATISTICAL REVIEW(S)

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:	21876 / N0002
Drug Name:	Diclegis
Indication(s):	Treatment of nausea and vomiting of pregnancy in patients who do not respond to conservative management
Applicant:	Duchesnay, Inc.
Date(s):	Submission Date: 06/08/2012
	PDUFA Due Date: 04/08/2013
Review Priority:	Standard
Biometrics Division:	Division of Biometrics III
Statistical Reviewer:	Kate Dwyer, Ph.D.
Concurring Reviewer:	Mahboob Sobhan, Ph.D., Team Leader
Medical Division:	Division of Reproductive and Urologic Drug Products
Clinical Team:	Theresa Van Der Vlugt, M.D., Medical Reviewer
	Shelley R. Slaughter, M.D., Team Leader
Project Manager:	George Lyght

Keywords: NDA review, clinical studies

Table of Contents

LIST OF TABLES

1 EXECUTIVE SUMMARY

The Applicant, Duchesnay, is seeking approval for Diclegis, a delayed release tablet containing 10 mg of doxylamine succinate and 10 mg of pyridoxine hydrochloride, for the treatment of nausea and vomiting of pregnancy in patients who do not respond to conservative management. To support this claim, one study (DIC-301) was submitted to this application.

Study DIC-301 was a double-blind, randomized, multi-center, placebo-controlled, parallel-group study conducted in the US in women aged 18 years or older with nausea and vomiting of pregnancy (NVP) and a Pregnancy Unique Quantification of Emesis (PUQE) score ≥ 6. The primary efficacy endpoint was the change in PUQE score from baseline to Day 15 ($\pm$ 1 day).

There were no statistical issues with this submission. Applicant's analysis based on ITT population and confirmed by FDA analysis, showed statistically significant improvement in PUQE score between the Diclegis and the placebo group. The treatment difference in the change from baseline was -0.73 (95% C.I: -1.25 to -0.22), favoring Diclegis over placebo. Sensitivity analyses using complete data and per protocol (PP) populations showed inconsistent efficacy results. There was a small treatment difference in PUQE score only in the PP population, but not in the subjects who completed the entire study and in the subjects who completed the entire study without any major protocol violations.

Secondary analyses of three individual components of PUQE showed a significant improvement in two components: number of vomits and number of retching. One other secondary endpoint of global assessment of well-being score also showed a significant treatment effect of Diclegis over placebo. However, the study was neither powered to test the above secondary hypotheses nor planned for any multiplicity adjustment. Therefore, secondary analyses were considered exploratory.

From a statistical perspective, the data submitted from study DIC-301 provided evidence of efficacy, albeit small one, in support of Diclegis in the treatment of pregnant women with NVP. However, the clinical applicability of this small treatment benefit over placebo is a clinical decision.

2 INTRODUCTION

2.1 Overview

The Applicant, Duchesnay, is seeking marketing approval for Diclegis, a delayed release tablet containing 10 mg of doxylamine succinate and 10 mg of pyridoxine hydrochloride, for the treatment of nausea and vomiting of pregnancy in patients who do not respond to conservative management.

The efficacy and safety of Diclegis was assessed in one phase 3 study. As shown in Table 1, study DIC-301 was a double-blind, randomized, multi-center, placebo-controlled, parallel group study of Diclegis for the treatment of pregnant women at least 18 years old with nausea and vomiting of pregnancy (NVP) and a Pregnancy Unique Quantification of Emesis (PUQE) score ≥ 6.

Table 1: Brief summary of the phase 3 studies for Diclegis

Study Number (No. of Sites / Country) Dates of Study Conduct	Subject Population	Primary Endpoints	Treatments	R (ITT)[1]	Design[2]
DIC-301 (6 / U.S.) Feb. 2008 to Jun. 2009	Women 18+ years of age with NVP and PUQE score ≥ 6	The change from baseline in PUQE score at Day 15 (± 1 day)	Diclegis Placebo **Total**	236 (235) 229 (224) **465 (459)**	DB, R, MC, PC

[1] R= Randomized, ITT = Intent-To-Treat
[2] DB = Double Blind, R = Randomized, MC = Multicenter, PC = Placebo Controlled

2.2 Data Sources

The study reports and the data sets were submitted electronically to the Electronic Document Room. The SAS data sets were complete and well documented.

The study reports are located at:

\\Cdsesub1\evsprod\NDA021876\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud

The datasets and programs for Study DIC-301 are located at:

\\Cdsesub1\evsprod\NDA021876\0000\m5\datasets\dic-301

\\Cdsesub1\evsprod\NDA021876\0014\m5\datasets\dic-301

3 STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Study Design and Endpoints

Study DIC-301 was a double-blind, randomized, multi-center, placebo-controlled study of Diclegis in the treatment of pregnant women at least 18 years old with NVP and PUQE score ≥ 6. The study was conducted in 6 US sites between February, 2008 and June 2009. Subjects were assigned in a 1:1 ratio to Diclegis and placebo, respectively.

Two tablets of study drug (Diclegis or placebo) were administered at bedtime on Day 1. If symptoms of nausea and vomiting persisted into the afternoon hours of Day 2, the subject was directed to take her usual dose of 2 tablets at bedtime and an additional tablet the next morning on Day 3. Based upon assessment in the clinic on Day 4 (± 1 day), the subject was directed to take an additional tablet mid-afternoon to control evening symptoms. The minimum assigned study medication was 2 tablets daily at bedtime for 14 days, increasing when indicated to the maximal dosage of 4 tablets per day according to the timing, duration, severity, and frequency of the symptoms experienced by the subject.

Subjects were scheduled to return to the clinic for evaluation in the morning prior to their morning dose on Day 4 (± 1 day), Day 8 (± 1 day), and then on Day 15 (± 1 day) for an end of study visit. Additionally, telephone contact was made at day 2, 6, 12, and 14 in order to assess subject diary information, adverse events (AEs), concomitant medications, and compliance with the study medication.

The primary objective of this study was to compare the efficacy of Diclegis to that of placebo in the treatment of nausea and vomiting of pregnancy (NVP) when administered orally daily for approximately 2 weeks.

The primary efficacy endpoint was the change from baseline in PUQE score at Day 15 (± 1 day). Change from baseline was calculated as post-baseline score minus baseline value. The PUQE incorporated the number of daily vomiting episodes, number of daily heaves, and length of daily nausea in hours, for an overall score of symptoms rated from 3 (no symptoms) to 15 (most severe). Baseline was defined as the PUQE score completed at the enrollment visit. For subjects who discontinued the study prematurely, last-observation-carried-forward (LOCF) approach for the subsequent visits was used for missing PUQE scores.

The secondary efficacy endpoints included the following:

a) Three components constituting the PUQE;
b) Global Assessment of Well-Being;
c) Number of tablets taken;
d) Time loss from household tasks and or employment;
e) Total number of visits and phone calls to health care providers;
f) Rates of hyperemesis gravidarum;
g) Compliance with study medication (0 = less than 28 tablets, 1 = 28 tablets, 2 = more than 28 tablets).

The exploratory endpoint was the relationship between levels of vitamin B6 (total and metabolites) and doxylamine and the severity of NVP symptoms (PUQE score).

For Global Assessment of Well Being, missing values were imputed using the last observation-carried-forward method for the subsequent visits for patients who discontinue the study prematurely. For other efficacy variables, analysis was based on the available data.

For time loss from household tasks and/or employment, and for number of visits and phone calls to health care providers, the numbers was adjusted to 15 days. Missing days were imputed with the period mean. The following examples were provided for illustration purposes:

1) A patient discontinued from the study on Day 5 and lost 4 hours from employment between Day 1 and discontinuation. Time loss from employment for this patient during the 15-day period is (4 hours out of 5 days)*15 days) = 12 hours.
2) A patient completed the study on Day 16 and lost 4 hours from employment between Day 1 and Day 16. Time loss from employment for this patient is (4 hours out of 16 days)*(15 days) = 3.75 hours.

The following four analyses populations were pre-specified in the protocol.

ITT-Safety: Any subject who took at least one dose of study medication during the study.

ITT-Efficacy: Any subject who took at least one dose of study medication and had at least one post-baseline PUQE measurement.

Complete Data: Any subject who (a) has recorded baseline PUQE score, (b) has recorded PUQE scores for at least 7 of the 14 expected daily diaries from the second day of the subject's maximal dose taken to Day 15 (± 1 day), and (c) absence of any major protocol violations including the violation of entry criteria.

Per protocol (PP): Any subject who (a) has a valid baseline assessment, (b) has recorded Day 15 (± 1 day) PUQE scores, (c) completed the study with between 80% - 120% of prescribed study medication applications, and (d) absence of any major protocol violations including the violation of entry criteria.

Reviewer's Comments: *In general completer population should include any subjects who completed the study without exclusion of protocol violators. Therefore, this review will not exclude protocol violators from the complete data population.*

Revised Completer Population: Any subject who (a) has recorded baseline PUQE score, (b) has recorded PUQE scores for at least 7 of the 14 expected daily diaries from the second day of the subject's maximal dose taken to Day 15 (± 1 day).

The primary efficacy analyses were conducted using ITT-E analysis population. Two additional populations, complete data and per protocol population, were also used as sensitivity analyses of primary efficacy analysis. This reviewer also conducted a sensitivity analysis using revised completer population, where protocol violators were not excluded.

Assuming the expected difference of 3 in PUQE scores between Diclegis® and placebo, a sample size of 64 per group was calculated to provide 90% power at alpha level of 0.001. A total of 280

patients (140 patients per treatment group) were enrolled in order to have 200 evaluable patients. With estimated dropout rate of 25% and a non-compliance rate of approximately 5%, this sample size was at least 4-fold larger than needed to show the intended clinical effect.

3.1.2 Statistical Methodologies

PUQE score was evaluated using an analysis of covariate (ANCOVA) model where change from baseline to Day 15 (± 1 day) was the response variable, the baseline PUQE score was the covariate, and the treatment group and study center were the fixed effects. The following ANCOVA assumptions were tested at 5% significance level unless otherwise noted: (1) normality of errors, (2) homogeneity of variances, and (3) equality of slopes among treatment groups at 10% significance level. If the assumptions were severely violated, a nonparametric approach (rank-based analysis of covariance method) was be used, stratifying by study center.

To show effectiveness for the treatment of NVP, the change in PUQE score from baseline to 15 days of post-treatment should be larger in the Diclegis arm than in the placebo arm in ITT-E population via LOCF.

For PUQE Score, sensitivity analyses was performed to examine the impact of missing data, and hence to demonstrate that study conclusions were invariant to assumptions, the particular model, and methods of handling missing data. Sensitivity analyses were also performed using both completer and per protocol populations.

In addition to composite PUQE, further analysis included each of the 3 components of PUQE: number of hours of nausea, number of times vomiting and number of retching episodes. Change in each of these components were compared between the two treatment groups using ANCOVA where change from baseline to Day 15 (± 1 day) was the response variable, the baseline value was the covariate, and the treatment group and study center are the fixed effects. If the assumptions were severely violated, a nonparametric approach (rank-based analysis of covariance method) was used, stratifying by study center.

Change in the global assessment of well being was compared between the two treatment groups using ANCOVA where change from baseline to Day 15 (± 1 day) via LOCF was the response variable, the baseline value was the covariate, and the treatment group and study center were the fixed effects. If the assumptions were severely violated, a nonparametric approach (rank-based analysis of covariance method) was used, stratifying by study center.

Number of tablets taken, total number of visits and phone calls to health care providers, and time loss from household tasks and/or employment was analyzed using an ANOVA model where period total was the response variable and the treatment and study center were the fixed effects. If the assumptions (normality of errors and homogeneity of variances) were severely violated, a nonparametric approach (rank-based analysis of variance method) was used.

Compliance with study medication and rates of hyperemesis gravidarum was examined using the Cochran-Mantel-Haenszel (CMH) row mean scores test controlling for study center.

All the secondary analyses were performed using ITT-E population.

3.1.3 Patient Disposition, Demographic and Baseline Characteristics

As shown in Table 2, a total of 280 women were enrolled and randomized in study DIC-301. More patients discontinued the in the placebo group than the Diclegis-treated group (35% vs. 20%). The most common reasons for premature discontinuation was withdrawal of consent (6.4% vs. 12.9%), followed by lost to follow up (5.0% vs. 13.6%) Diclegis compared to placebo group, respectively.

Table 2: Subject Disposition: Study DIC-301

Category	Diclegis		Placebo		Total	
Randomized	140		140		280	
Treated (ITT-S)	133	95.0%	128	91.4%	261	93.2%
ITT-E	131	93.6%	125	89.3%	256	91.4%
PP	103	73.6%	79	56.4%	182	65.0%
Complete Data	101	72.1%	75	53.6%	176	62.9%
Completed Study[1]	112	80.0%	91	65.0%	203	72.5%
Discontinued the Study	28	20.0%	49	35.0%	77	27.5%
Reason for Discontinued						
Adverse event	5	3.6%	5	3.6%	10	3.6%
Withdrew consent	9	6.4%	18	12.9%	27	9.6%
Treatment Failure	2	1.4%	5	3.6%	7	2.5%
Lost to Follow-up	7	5.0%	19	13.6%	26	9.3%
Others	5	3.6%	2	1.4%	7	2.5%

* Percentage based on the total number of treated subjects within each corresponding treatment group.
[1] Reviewer's defined Revised Completer population consists of any subject who completed the study.
(Source: Reviewer's Analysis and Clinical Study DIC-301 Report; Table 10.1, page 41)

Demographic and baseline characteristics all treated subjects are shown in Table 3. The mean age was 25.5 years (range 18 – 42). The majority of patients were Caucasian (59.4%) and black or African American (37.9%). The mean BMI of the patients was 29.3 (range 11.6 – 116.8). The average gestational age at enrollment was 9.3 weeks (range 7 - 14 weeks). The mean PUQE score at enrollment was similar in both groups.

Table 3: Demographic and Baseline Characteristics (Categorical Variables): Study DIC-301, ITT-S

Variables	Diclegis N=133	Placebo N=128
Maternal Age (Yr) (SD)	25.9 ± 6.0 Median: 25.0 Min-Max: 18-45	25.0 ± 5.6 Median: 23.5 Min-Max: 18-42
Race (%)		
White	60.2	58.6
African American	37.6	38.3
Other	2.2	3.1
BMI (kg/m^2) (SD)	28.9 ± 7.6 Median: 28.0 Min-Max: 16.7-53.2	29.8 ± 11.1 Median: 26.9 Min-Max: 11.6-116.8
BMI Group[1]		
Underweight	5 (3.8%)	4 (3.1%)
Normal	39 (29.3%)	38 (29.7%)
Overweight	31 (23.3%)	41 (32.0%)
Obese	57 (42.9%)	44 (34.4%)
Gestational Age at Enrollment (Weeks) (SD)	9.3 ± 1.9 Median: 9 Min-Max: 7-13	8.8 ± 2.1 Median: 9 Min-Max: 7-14
PUQE score at Enrollment (SD)	9.0 ± 2.1 Median: 9 Min-Max: 6-15	8.8 ± 2.1 Median: 8 Min-Max: 5-15

[1]BMI group: Underweight below 18.5, Normal 18.5-24.9, Overweight 25.0-29.9, Obese 30.0 and above.
(Source: Reviewer's Analysis and Adapted from Clinical Study DIC-301 Report; Table 10.2 & 10.3, page 43-44)

3.1.4 Results and Conclusion

The primary efficacy endpoint was the change from baseline in PUQE score at Day 15 (± 1 day). Subjects who discontinued the study prematurely, their missing PUQE scores were imputed by LOCF approach. The results of the primary efficacy analysis are presented in Table 4. The results showed that there was statistically significantly greater improvement in the Diclegis group than in the placebo group.

Table 4. Primary Endpoint Analysis: Change from Baseline to Day 15 (± 1 day) in PUQE Score (ITT-E Population)

Data/ Category	Statistics	Treatment Group Diclectin® (N = 131)	Treatment Group Placebo (N = 125)	*P* value for comparison
ITT-E via LOCF				
Baseline	N	131	125	
	Mean ± SD	9.0 ± 2.1	8.8 ± 2.1	
	Median	9.0	8.0	
	Min, Max	6, 15	6, 15	
Day 15 (± 1 day)	N	131	125	
	Mean ± SD	4.2 ± 1.9	4.9 ± 2.3	
	Median	3.0	4.0	
	Min, Max	3, 11	3, 12	
Change from Baseline	N	131	125	
	Mean ± SD	-4.8 ± 2.7	-3.9 ± 2.6	0.006[1]
	Median	-5.0	-4.0	
	Min, Max	-11, 3	-11, 2	

(Source: Clinical Study DIC-301 Report; Table 11.1, page 47)

The Application did not provide a point estimate of the treatment difference between Diclegis and placebo and its 95% confidence interval for PUQE score and its three components. An information request (IR) was sent to the Applicant on November 8, 2012. Subsequently, the Applicant submitted additional efficacy analyses in response to our IR on December 5, 2012. The Applicant and independent analyses conducted by this review showed a treatment difference in change from baseline -0.73 (95% C.I: -1.25 to -0.22) in PUQE score.

Two of three pre-specified important components of PUQE: numbers of times vomiting and number of retching episodes, were considered important secondary endpoints for this indication. Based on ITT-E population, number of vomiting retching showed a significant improvement, but there was no difference in hours of nausea. This reviewer conducted additional sensitivity analyses using revised completer population for PUQE and its three components. The results in Table 7 showed that there was no treatment difference in PUQE score and only the third component: number of retching episodes had statistically significant improvement of 0.18.

Table 5. Applicant Analyses: Change from Baseline to Day 15 (± 1 day) in PUQE Score and its Three Components (ITT-E Population via LOCF)

Population	Diclegis (n=131)	Placebo (n=125)	Diclegis vs. Placebo	
	Change from Baseline	Change from Baseline	Diff. (95% C.I.)	p-value
PUQE Score	-4.67	-3.94	-0.73 (-1.24,-0.22)	0.006
Hours of Nausea	-2.35	-2.13	-0.21 (-0.49, 0.06)	0.125
Number of Vomited	-0.95	-0.72	-0.23 (-0.39,-0.06)	0.008
Number of Retching	-1.37	-1.1	-0.28 (-0.46,-0.09)	0.003

(Source: Adapted from Response to FDA Request for Information: Clinical Statistical Analysis)

Table 6. FDA Analyses: Change from Baseline to Day 15 (± 1 day) in PUQE Score and its Three Components (ITT-E Population via LOCF)

Population	Diclegis (n=131)	Placebo (n=125)	Diclegis vs. Placebo	
	Change from Baseline	Change from Baseline	Diff. (95% C.I.)	p-value
PUQE Score	-4.67	-3.94	-0.73 (-1.25,-0.22)	0.006
Hours of Nausea	-2.35	-2.13	-0.21 (-0.49, 0.06)	0.126
Number of Vomited	-0.95	-0.72	-0.22 (-0.39,-0.06)	0.008
Number of Retching	-1.37	-1.1	-0.28 (-0.46,-0.09)	0.004

(Source: Reviewer's Analyses)

Table 7. FDA Sensitivity Analyses: Change from Baseline to Day 15 (± 1 day) in PUQE Score and its Three Components (Revised Completer Population), LOCF

Population	Diclegis (n=112)	Placebo (n=91)	Diclegis vs. Placebo	
	Change from Baseline	Change from Baseline	Diff. (95% C.I.)	p-value
PUQE Score	-5.06	-4.72	-0.34 (-0.79, 0.11)	0.140
Hours of Nausea	-2.53	-2.47	-0.06 (-0.32, 0.20)	0.637
Number of Vomited	-1.03	-0.95	-0.08 (-0.19, 0.03)	0.156
Number of Retching	-1.49	-1.31	-0.18 (-0.35,-0.01)	0.035

(Source: Reviewer's Analyses)

Results of sensitivity analyses using Per Protocol and complete data population showed a significant treatment difference of 0.49 only in PP population, but not in complete population.

Table 8. Additional Sensitivity Analyses: Change from Baseline to Day 15 (± 1 day) in PUQE Score

Population	Diclectin		Placebo		Diclectin vs. Placebo	
	n	Change from Baseline	n	Change from Baseline	Diff. (95% C.I.)	p-value
Per Protocol	103	-5.13	79	-4.64	-0.49 (-0.96,-0.01)	0.044
Complete Data	101	-4.93	75	-4.57	-0.36 (-0.86, 0.13)	0.148

(Source: Reviewer's Analyses)

Additional sensitivity analyses using mixed model repeated measures (MMRM) model (using daily measure from day 1 to day 14 without LOCF imputation) was later submitted by the Applicant upon request. These analyses were also confirmed by FDA analyses to explore the impact of imputation for missing values. The results showed that there were significant treatment improvements in PUQE score and three of its components.

Among other secondary endpoints, this reviewer confirmed that only global assessment of well-being score showed a significant treatment increase over placebo. There was a small treatment improvement of 0.69 (95% C.I: 0.18 to 1.19) in global assessment of well-being score out of total 10 score. There were no treatment differences in secondary endpoints including number of tablets taken, time loss from household tasks and or employment, total number of visits and phone calls to health care providers, rates of hyperemesis gravidarum, and compliance with study medication.

3.2 Evaluation of Safety

Statistical evaluation of safety is not necessary for this review. Safety information is found in the clinical "review of safety" section.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

The efficacy results were based on women aged 18 up to 45 only so the subgroup analyses by gender and age are not necessary for this indication. The phase 3 study was conducted in the US so no by-region analysis is necessary.

The Applicant did not submit any subgroup analyses. I conducted subgroup analyses by race. As presented in Table 9, black women had almost twice treatment improvement of Diclegis over placebo than that of white woman.

Table 9: Primary Endpoint Analysis: Change from Baseline to Day 15 (± 1 day) in PUQE Score by Race (ITT-E Population)

Population	Diclectin		Placebo		Diclectin vs. Placebo	
	n	Change from Baseline	n	Change from Baseline	Diff. (95% C.I.)	p-value
White	80	-4.59	73	-4.01	-0.58 (-1.28, 0.11)	0.101
African American	49	-4.96	48	-3.89	-1.06 (-1.88,-0.25)	0.010

(Source: Reviewer's Analyses)

4.2 Other Subgroup Populations

No other efficacy subgroup analysis was evaluated.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

There were no statistical issues with this submission. Efficacy was evaluated by measuring the change from baseline in PUQE score at Day 15 (± 1 day). Results from the Applicant and confirmed by FDA showed a small improvement in PUQUE score of -0.73 (95% C.I: -1.25 to -0.22) in pregnant women with NVP. Sensitivity analyses using other populations showed inconsistent efficacy. There was a small treatment difference in PUQE score only in the PP population, but not in the subjects who completed the entire study and in the subjects who completed the entire study without any major protocol violations. Furthermore, in three components of the PUQE composite score, there was a significant improvement in the number of vomiting and retching, but not in the number of hours of nausea. Among all other secondary endpoints, only global assessment of well-being score showed a significant treatment effect of Diclegis over placebo in pregnant women with NVP.

5.2 Conclusions and Recommendations

From a statistical perspective, the data submitted from the study DIC-301 provided some evidence that Diclegis was effective in the treatment of pregnant women with NVP. However, clinical significance of such a small treatment effect and approvability decision is a clinical call.

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

/s/

KATE L DWYER
03/09/2013

MAHBOOB SOBHAN
03/11/2013