

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

21-992

STATISTICAL REVIEW(S)



DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH

STATISTICAL REVIEW AND EVALUATION

Clinical Studies

NDA/Serial Number: 21-992 (N000)
Drug Name: Pristiq™ (desvenlafaxine) Extended-Release Tablets
Indication: MDD
Applicant: Wyeth
Dates: Date of Document: 8/29/2007
PDUFA Due Date: 3/1/2008
Review Priority: Priority
Biometrics Division: Biometrics I, HFD-710
Statistical Reviewer: Yeh-Fong Chen, Ph.D.
Concurring Reviewers: Peiling Yang, Ph.D.
James Hung, Ph.D.
Medical Division: Division of Psychiatry Products, HFD-130
Clinical Team: Clinical Reviewer: Robert Levin, M.D.
Clinical Team Leader: Gwen Zornberg, M.D.
Project Manager: Renmeet Gujral, Pharm. D., HFD-130

Table of Contents

1. EXECUTIVE SUMMARY	3
1.1 CONCLUSIONS AND RECOMMENDATIONS	3
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES	3
1.3 STATISTICAL ISSUES AND FINDINGS	3
2. INTRODUCTION	4
2.1 OVERVIEW	4
2.2 DATA SOURCES	5
3. STATISTICAL EVALUATION	5
3.1 EVALUATION OF EFFICACY.....	5
3.1.1 Description of Studies 332 and 333.....	5
3.1.1.1 Study Objectives	5
3.1.1.2 Study Design.....	6
3.1.1.3 Efficacy Variables and Analyses	6
3.1.2 Efficacy Results for Study 332	7
3.1.2.1 Patient Disposition and Baseline Demographic Characteristics.....	7
3.1.2.2 Sponsor's Efficacy Results for Primary Endpoint.....	8
3.1.2.3 Sponsor's Efficacy Results for Key Secondary Endpoint.....	9
3.1.2.4 Sponsor's Efficacy Results for Other Secondary Endpoints	9
3.1.2.5 Statistical Reviewer's Findings and Comments	10
3.1.3 Efficacy Results for Study 333	12
3.1.3.1 Patient Dispositions and Baseline Demographic Characteristics	13
3.1.3.2 Sponsor's Efficacy Results for Primary Endpoint.....	14
3.1.3.3 Sponsor's Efficacy Results for Key Secondary Endpoint	15
3.1.3.4 Sponsor's Efficacy Results for Other Secondary Endpoints	16
3.1.3.5 Statistical Reviewer's Findings and Comments	16
3.2 EVALUATION OF SAFETY.....	16
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	17
4.1 GENDER, RACE AND AGE	17
4.1.1 For Study 332	17
4.1.2 For Study 333	18
4.2 OTHER SPECIAL/SUBGROUP POPULATIONS.....	19
5. SUMMARY AND CONCLUSIONS	20
5.2 CONCLUSIONS AND RECOMMENDATIONS.....	20

1. EXECUTIVE SUMMARY

1.1 CONCLUSIONS AND RECOMMENDATIONS

Among the two short term studies that the sponsor submitted to evaluate the efficacy of DVS SR 50-mg and DVS SR 100-mg, only Study 333 showed strong evidence of efficacy for both doses. Although the sponsor's analysis results showed that Study 332 supported the efficacy of DVS SR 50-mg, this reviewer had concerns about the cleanness of the data in the sense that some patients have unexpected data. When the unexpected data were removed, the statistical significance of the overall study and DVS SR 50-mg in the primary analysis appears to be diminished. In addition, for Study 332, this reviewer's MMRM analysis results showed larger p-values than the sponsor's although p-values were still less than the nominal significance level 0.05 for both dose levels.

In both studies, the effects of two doses were numerically similar regardless of statistical significance. However, results in Studies 332 and 333 were in favor of the 50 mg dose and 100 mg dose, respectively.

1.2 BRIEF OVERVIEW OF CLINICAL STUDIES

This submission is to complete the sponsor's response to the FDA approvable letter dated 1/22/2007, where the sponsor was advised to conduct short term trials to explore the dose response of the drug for effectiveness, especially for the doses lower than 100 to 400 mg/day, and also a long-term trial for the drug's maintenance efficacy.

In responding to the letter, the sponsor submitted three new efficacy studies. Two studies had the acute phase of eight-weeks (Studies 332 and 333) and one was a randomized withdrawal trial with a 12 week open-label phase and 6-month double-blind phase (Study 302). Since the long-term study 302 used much higher doses (DVS SR 200 or 400 mg/day) than two short-term studies (DVS SR 50 mg and 100 mg), which is not reasonable, the medical division decided not to accept Study 302 and that there is no need to perform thorough review for Study 302. Therefore, only the two fixed dosed 8 week studies (Studies 332 and 333) are evaluated and reported in this statistical review.

Studies 332 and 333 had an identical design except that Study 332 was conducted in the United States and Study 333 was conducted in countries of Europe and South Africa. The primary endpoint for these studies was the adjusted mean change from baseline on the HAM-D₁₇ total score at the final on-therapy evaluation. The primary analysis was the analysis of covariance (ANCOVA) with treatment and site as factors and baseline scores as the covariate.

1.3 STATISTICAL ISSUES AND FINDINGS

This reviewer found that for both Studies 332 and 333, a small percentage of patients (around 5% to 10%) had the unplanned Weeks 5 and 7 data. Although the percentages are small, there is still a possibility to bias the analysis results. This reviewer thus performed the re-analyses by removing the unexpected data. The summary of this reviewer's analysis results for the final on-therapy evaluation and the MMRM are shown in Table 2.1.1. Note that the OC analysis results are not affected by this data removal.

According to the reviewer's analysis results, Study 332 appears to be negative but Study 333 still positive although Study 332's overall p-value was very close to 0.05. Based on the sponsor pre-specified closed testing method for multiplicity, none of doses in Study 332 should be further tested after knowing that the overall p-value >0.05. On the contrary, both 50 mg and 100 mg DVS SR showed statistical significant results in comparison with placebo for Study 333.

Table 2.1.1 Summary of Reviewer's Analysis Results for Primary Endpoint, Change from Baseline on the HAM-D₁₇ Total Score

Analysis		Study 332			Study 333		
		Placebo	DVS SR 50 mg	DVS SR 100 mg	Placebo	DVS SR 50 mg	DVS SR 100 mg
Final On- Therapy Evaluation	N	150	150	147	161	164	158
	LSM ¹	-9.57	-11.49	-10.96	-10.78	-13.16	-13.65
	P-Value ²		0.02	0.09		0.004	<0.001
	P-Value ³	0.054			0.001		
MMRM Analysis	N	115	104	102	138	145	126
	LSM ¹	-9.81	-12.46	-11.73	-11.47	-14.25	-14.90
	P-Value ²		0.003	0.03		<0.001	<0.001
	P-Value ³	0.02			0.006		

¹ LSM stands for least square mean

² These P-Values are for the comparisons with placebo

³ This P-Value is for the overall comparison

2. INTRODUCTION

2.1 OVERVIEW

This submission is to complete the sponsor's response to the FDA approvable letter dated 1/22/2007, where the sponsor was advised to conduct short term trials to explore the dose response of the drug for effectiveness, especially for the doses lower than 100 to 400 mg/day, and also a long-term trial for the drug's maintenance efficacy.

This DVS SR submission includes the final clinical study reports (CSRs) for 4 phase I Studies (studies 198, 401, 900, and 902) and 4 phase 3 MDD studies (studies 302, 303, 332 and 333). Among these four phase III MDD studies, Studies 332 and 333 had the acute phase of eight-weeks and Study 302 was a randomized withdrawal trial with a 12 week open-label phase and 6-month double-blind phase. Study 303 was designed as an open label phase 3 study, so the statistical review was not performed. On the other hand, since the long-term study 302 used much higher doses (DVS SR 200 or 400 mg/day) than two short-term studies (DVS SR 50 mg and 100 mg), the medical division decided not to accept Study 302 and that there is no need to perform thorough statistical review for Study 302. Therefore, only the two fixed dosed 8 week studies (Studies 332 and 333) are evaluated and reported in this statistical review.

Studies 332 and 333 had an identical design except that Study 332 was conducted in the United States and Study 333 was conducted in countries of Europe and South Africa. The primary endpoint for these studies was the adjusted mean change from baseline on the HAM-D₁₇ total score at the final on-therapy evaluation. The primary analysis is the analysis of covariance (ANCOVA) with treatment and site as factors and baseline scores as the covariate.

Table 2.1.2 summarizes the sponsor's analysis results for the primary endpoint for both studies based on the primary analysis, i.e., final on therapy evaluation, based on the secondary analyses, observed cases analysis (i.e., OC) and the Mixed Effect Model of Repeated Measure analysis (i.e., MMRM). According to the sponsor's analysis results, Study 332 showed the efficacy of the DVS SR 50mg and as well as Study 333 although only Study 333 showed the efficacy of the DVS SR 100 mg.

Table 2.1.2 Summary of Sponsor's Analysis Results for Primary Endpoint, Change from Baseline on the HAM-D₁₇ Total Score

Analysis		Study 332			Study 333		
		Placebo	DVS SR 50 mg	DVS SR 100 mg	Placebo	DVS SR 50 mg	DVS SR 100 mg
Final On- Therapy Evaluation	N	150	150	147	161	164	158
	LSM ¹	-9.53	-11.5	-11.0	-10.7	-13.2	-13.7
	P-Value ²		0.018	0.065		0.002	<0.001
	P-Value ³		0.046			<0.001	
OC Analysis	N	115	104	102	138	145	126
	LSM ¹	-10.0	-12.1	-11.9	-11.6	-14.7	-15.2
	P-Value ²		0.026	0.047		<0.001	<0.001
	P-Value ³		0.049			<0.001	
MMRM Analysis	N	115	104	102	138	145	126
	LSM ¹	-9.86	-12.40	-11.88	-11.49	-14.37	-14.91
	P-Value ²		<0.001	0.006		<0.001	<0.001
	P-Value ³		Not Available			Not Available	

¹ LSM stands for least square mean

² These P-Values are for the comparisons with placebo

³ This P-Value is for the overall comparison

2.2 DATA SOURCES

This submission is stored in the following directory of the CDER's EDR:
\\Cdsesub1\evsprod\NDA021992\0033. The sponsor's response to FDA's questions regarding data and MMRM analyses is stored in \\Cdsesub1\evsprod\NDA021992\0037.

3. STATISTICAL EVALUATION

3.1 EVALUATION OF EFFICACY

The following description is based on the sponsor's clinical study report. Any discrepancy between the study report and study protocol will be discussed in the section of statistical reviewer's comments.

3.1.1 Description of Studies 332 and 333

3.1.1.1 Study Objectives

The primary objective was to compare the antidepressant efficacy, safety, and tolerability of DVS SR in subjects receiving daily doses of 50 mg or 100 mg of DVS SR versus subjects receiving placebo. Additional objectives included testing both general and functional quality-of-life outcomes and satisfaction with therapy reported by the subject.

3.1.1.2 Study Design

This was a phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group study in adult outpatients with major depressive disorder (MDD).

After a screening period of 6 to 14 days, eligible subjects were treated for approximately 8 weeks plus 1 additional week for tapering the study drug. Subjects assigned to the 100-mg dose group had their dose titrated to 50 mg/day during week 1. Subjects returned for a post-study/follow-up visit (day 70) approximately 7 days after discontinuing use of the study drug.

3.1.1.3 Efficacy Variables and Analyses

The primary efficacy variable was the HAM-D₁₇ total score. The key secondary efficacy variable was the CGI-I score, which was rated from 1 (very much improved) to 7 (very much worse).

Other secondary variables included:

- MARDS total score
- CGI-S score
- Remission rate (percentage of subjects with HAM-D₁₇ scores of ≤ 7)
- Response rates for the HAM-D₁₇, MADRS, and CGI-I
- VAS-PI overall and component scores
- HAM-D₆ (Bech version: HAM-D, items 1, 2, 7, 8, 10 and 13) total score
- Covi Anxiety Scale total score

For all primary and secondary efficacy variables, the final on-therapy evaluation was the primary endpoint. All efficacy measures such as the HAM-D₁₇ and MADRS total score were evaluated similarly by using analysis of covariance (ANCOVA) on changes from baseline with treatment and site as the factors and baseline scores as the covariate. An exception to this was the CGI-I. CGI-I scores were analyzed by using the Cochran-Mantel-Haenszel test as the primary analysis and by using ANOVA with treatment and site as factors as a secondary analysis. These analyses were done at each evaluation period by using the LOCF technique and the observed-case analysis.

For the primary efficacy variable, closed testing procedures were performed to compare the 2 doses (50 and 100mg/day) of DVS SR with placebo. A general liner model with multiple contrast statements was used to calculate F-statistics for the global null hypothesis and all intersection hypotheses. The closure principle was used to determine which hypothesis should have been retained or rejected at $\alpha=0.05$.

If a significant difference was detected for one or both doses of DVS SR, then a sequential testing method was applied to that dose(s) as follows: For one or both DVS SR dose group(s), if a significant difference from placebo on the primary efficacy variable was noted based on the closed testing procedure, the key secondary efficacy variable was tested at the 0.05 level to compare the DVS dose(s) with placebo.

3.1.2 Efficacy Results for Study 332

3.1.2.1 Patient Disposition and Baseline Demographic Characteristics

A total of 703 subjects were screened for participation in this study; 229 were screen failures and 474 were randomly assigned to treatment: 159 were assigned to receive placebo, 158 were assigned to receive DVS SR 50 mg, and 157 were assigned to receive DVS SR 100 mg. The ITT population included 447 subjects. Table 3.1.2.1 summarizes the number of subjects who discontinued treatment by the reason for withdrawal in each treatment group.

Table 3.1.2.1 Number (%) of Subjects Who Discontinued During the On-Therapy Period by Reason for Withdrawal for Study 332

Reason	Placebo (n=152)	DVS SR 50 mg (n=151)	DVS SR 100 mg (n=148)
Total	25 (16.4)	34 (22.5)	31 (20.9)
Adverse event	4 (2.6)	5 (3.3)	11 (7.4)
Failed to return	6 (3.9)	15 (9.9)	11 (7.4)
Investigator request	0	1 (0.7)	0
Other event	3 (2.0)	2 (1.3)	3 (2.0)
Protocol violation	1 (0.7)	2 (1.3)	0
Subject request unrelated to study	6 (3.9)	9 (6.0)	5 (3.4)
Unsatisfactory response efficacy	5 (3.3)	0	1 (0.7)

Source: Sponsor's Table 8-3 of CSR.

The ITT patients' demographic and baseline characteristics are shown in Table 3.1.2.2. As shown in the table, there was a significant difference ($p=0.046$) between the treatment groups for the baseline characteristic weight. A pair-wise analysis showed a significant difference only between the placebo and DVS SR 50-mg groups. The sponsor made a note that the difference in baseline weight between the two treatment groups, however, would not be expected to have any impact on the efficacy results.

Table 3.1.2.2 Demographic and Baseline Characteristics for ITT Population for Study 332

Characteristic	Placebo (n = 150)	DVS SR 50 mg (n = 150)	DVS SR 100 mg (n = 147)
Ethnic Origin, n (%)			
Hispanic or Latino	13 (8.67)	18 (12.00)	17 (11.56)
Non-Hispanic and Non-Latino	137 (91.33)	132 (88.00)	130 (88.44)
Height (cm)	(n = 149)	(n = 149)	(n = 147)
Mean	167.90	168.41	170.08
SD	8.60	9.45	10.52
Min, max	151.10, 198.10	144.80, 199.40	149.90, 213.40
Weight (kg)			
Mean ^a	79.90	85.69	82.92
SD	19.82	21.74	18.66
Min, max	49.50, 157.50	42.20, 152.10	49.50, 143.50

Baseline CGI-S score			
Mean	4.31	4.32	4.39
SD	0.48	0.48	0.54
Min, max	4.00, 6.00	4.00, 6.00	4.00, 6.00
Median	4.00	4.00	4.00
Baseline HAM-D ₁₇ total score			
Mean	23.02	23.37	23.35
SD	2.59	2.64	2.61
Min, max	20.00, 30.00	20.00, 32.00	20.00, 32.00
Median	23.00	23.00	23.00
Baseline CGI-S, n (%)			
(4) moderately ill	104 (69.3)	103 (68.7)	93 (63.3)
(5) markedly ill	45 (30.0)	46 (30.7)	50 (34.0)
(6) severely ill	1 (0.7)	1 (0.7)	4 (2.7)

a. Significant difference (p=0.046) between the placebo and DVS SR 50-mg groups based on 1-way analysis of variance with treatment as factor.

Source: Sponsor's Table 8-8 of CSR.

3.1.2.2 Sponsor's Efficacy Results for Primary Endpoint

The Sponsor's efficacy results for the primary endpoint are shown in Table 3.1.2.3. As seen in the table, for the HAM-D₁₇ total score at the final on-therapy evaluation, the adjusted mean change from baseline was statistically significantly greater for the DVS SR 50-mg group (-11.5) in comparison with the placebo group (-9.53). The observed-cases analysis for HAM-D₁₇ total scores for the DVS SR 50-mg group showed similar results. For the DVS SR 100-mg group at the final on-therapy evaluation, there was no statistically significant difference in the adjusted mean change from baseline compared with the placebo group, nor were there statistically significant differences between these two treatment groups at any other evaluation.

Table 3.1.2.3 Sponsor's Analysis Results for the HAM-D₁₇ Total Scores (by ANCOVA) for Study 332

Week of Therapy	Therapy Group	n	Raw Mean Score	Adj Change From Baseline	Std Error	Adj Means (95% CI)	Diff Adj Mean (95% CI) vs Pbo.	Overall p-Value	p-Value vs DVS SR 50 mg	p-Value vs DVS SR 100 mg
Baseline	Placebo	150	23.0			23.2 (23.2, 23.2)				
	DVS SR 50 mg	150	23.4			23.2 (23.2, 23.2)				
	DVS SR 100 mg	147	23.4			23.2 (23.2, 23.2)				
Week 1	Placebo	147	19.2	-3.81	0.37	19.4 (18.7, 20.2)		0.416	0.331	0.767
	DVS SR 50 mg	147	19.1	-4.31	0.37	18.9 (18.2, 19.7)	0.5 (-0.5, 1.5)			0.206
	DVS SR 100 mg	143	19.7	-3.65	0.38	19.6 (18.9, 20.3)	-0.2 (-1.2, 0.9)			
Week 2	Placebo	149	17.6	-5.45	0.43	17.8 (17.0, 18.6)		0.368	0.205	0.240
	DVS SR 50 mg	150	17.2	-6.20	0.42	17.0 (16.2, 17.9)	0.8 (-0.4, 1.9)			0.931
	DVS SR 100 mg	146	17.2	-6.15	0.43	17.1 (16.3, 17.9)	0.7 (-0.5, 1.9)			
Week 3	Placebo	149	15.9	-7.08	0.48	16.2 (15.2, 17.1)		0.344	0.147	0.387
	DVS SR 50 mg	150	15.3	-8.06	0.48	15.2 (14.2, 16.1)	1.0 (-0.3, 2.3)			0.559
	DVS SR 100 mg	147	15.7	-7.66	0.48	15.6 (14.6, 16.5)	0.6 (-0.7, 1.9)			
Week 4	Placebo	150	15.5	-7.61	0.50	15.6 (14.7, 16.6)		0.065	0.019	0.269
	DVS SR 50 mg	150	14.1	-9.25	0.50	14.0 (13.0, 15.0)	1.6 (0.3, 3.0)			0.219
	DVS SR 100 mg	147	15.0	-8.39	0.50	14.9 (13.9, 15.8)	0.8 (-0.6, 2.2)			
Week 6	Placebo	150	14.3	-8.86	0.53	14.4 (13.4, 15.4)		0.039	0.013	0.080
	DVS SR 50 mg	150	12.7	-10.7	0.53	12.5 (11.5, 13.6)	1.8 (0.4, 3.3)			0.469
	DVS SR 100 mg	147	13.2	-10.2	0.54	13.1 (12.0, 14.1)	1.3 (-0.2, 2.8)			
Week 8	Placebo	150	13.9	-9.34	0.58	13.9 (12.8, 15.0)		0.020	0.006	0.053
	DVS SR 50 mg	150	11.9	-11.5	0.58	11.7 (10.6, 12.8)	2.2 (0.6, 3.8)			0.430
	DVS SR 100 mg	147	12.6	-10.9	0.58	12.3 (11.2, 13.5)	1.6 (-0.0, 3.2)			
Final	Placebo	150	13.7	-9.53	0.58	13.7 (12.6, 14.9)		0.046	0.018	0.065
	DVS SR 50 mg	150	12.0	-11.5	0.58	11.8 (10.6, 12.9)	1.9 (0.3, 3.5)			0.604
	DVS SR 100 mg	147	12.4	-11.0	0.59	12.2 (11.1, 13.4)	1.5 (-0.1, 3.1)			

Adj = adjusted; ANCOVA = analysis of covariance; HAM-D₁₇ = Hamilton Rating Scale for Depression, 17-item; ITT = intent-to-treat; LOCF = last-observation-carried-forward; FOT = final on-therapy.

Analysis	Week of Therapy	Therapy Group	Raw Mean Score	Adj Change From Baseline		Adj Means (95% CI)	Diff Adj Mean (95% CI) vs Pbo.	Overall p-Value	p-Value vs DVS SR 50 mg	p-Value vs DVS SR 100 mg
				n	Std Error					
Observed Cases	Baseline	Placebo	150	23.0			23.2 (23.2, 23.2)			
		DVS SR 50 mg	150	23.4			23.2 (23.2, 23.2)			
		DVS SR 100 mg	147	23.4			23.2 (23.2, 23.2)			
	Week 1	Placebo	147	19.2	-3.81	0.37	19.4 (18.7, 20.2)		0.416	0.331
		DVS SR 50 mg	147	19.1	-4.31	0.37	18.9 (18.2, 19.7)	0.5 (-0.5, 1.5)		
		DVS SR 100 mg	143	19.7	-3.65	0.38	19.6 (18.9, 20.3)	-0.2 (-1.2, 0.9)		0.206
	Week 2	Placebo	144	17.6	-5.32	0.44	17.9 (17.0, 18.7)		0.166	0.082
		DVS SR 50 mg	141	17.0	-6.39	0.44	16.8 (15.9, 17.7)	1.1 (-0.1, 2.3)		0.130
		DVS SR 100 mg	134	17.0	-6.26	0.45	16.9 (16.1, 17.8)	0.9 (-0.3, 2.2)		0.841
	Week 3	Placebo	138	15.9	-7.09	0.50	16.2 (15.2, 17.1)		0.143	0.059
		DVS SR 50 mg	129	14.9	-8.44	0.52	14.8 (13.8, 15.8)	1.4 (-0.0, 2.7)		0.626
		DVS SR 100 mg	130	15.3	-8.09	0.51	15.2 (14.1, 16.2)	1.0 (-0.4, 2.4)		
	Week 4	Placebo	132	14.9	-7.94	0.52	15.3 (14.2, 16.3)		0.105	0.037
		DVS SR 50 mg	128	13.8	-9.49	0.53	13.7 (12.7, 14.7)	1.6 (0.1, 3.0)		0.482
		DVS SR 100 mg	126	14.4	-8.97	0.53	14.2 (13.2, 15.3)	1.0 (-0.4, 2.5)		
	Week 6	Placebo	124	13.7	-9.49	0.56	13.7 (12.7, 14.8)		0.033	0.015
		DVS SR 50 mg	106	11.7	-11.5	0.60	11.8 (10.6, 12.9)	2.0 (0.4, 3.6)		0.046
		DVS SR 100 mg	112	12.1	-11.1	0.59	12.1 (11.0, 13.3)	1.6 (0.0, 3.2)		0.647
	Week 8	Placebo	115	13.1	-10.0	0.65	13.2 (11.9, 14.4)		0.049	0.026
		DVS SR 50 mg	104	11.0	-12.1	0.69	11.1 (9.7, 12.4)	2.1 (0.3, 3.9)		0.047
		DVS SR 100 mg	102	11.3	-11.9	0.69	11.3 (9.9, 12.6)	1.9 (0.0, 3.7)		0.824
	Week >8	Placebo	15	12.8	--	--	--		--	--
		DVS SR 50 mg	13	10.2	--	--	--		--	--
		DVS SR 100 mg	15	9.9	--	--	--		--	--

Adj = adjusted; ANCOVA = analysis of covariance; HAM-D₁₇ = Hamilton Rating Scale for Depression, 17-item; ITT = intent-to-treat; LOCF = last-observation-carried-forward; FOT = final on-therapy.

Source: Table 9-1 of CSR.

3.1.2.3 Sponsor's Efficacy Results for Key Secondary Endpoint

The key secondary efficacy variable was the CGI-I score. CGI-I scores are rated from 1 (very much improved) to 7 (very much worse). Table 3.1.2.4 shows the results of the key secondary efficacy endpoint, the percentage of subjects with each CGI-I score at the final on-therapy evaluation. Based on the sponsor's analysis results, at the final on-therapy evaluation, the CGI-I scores did not differ significantly for the DVS SR 50-mg group and DVS SR 100-mg group in comparison with the placebo group.

Table 3.1.2.4 Sponsor's Analysis Results for Secondary Endpoint CGI-I based on the LOCF data for Study 332

Week of Therapy	Therapy Group	n	1 n (%)	2 n (%)	3 n (%)	4 n (%)	5 n (%)	p-value vs. placebo
Week 8	Placebo	150	31 (20.7)	37 (24.7)	42 (28.0)	38 (25.3)	2 (1.3)	--
	DVS SR 50 mg	150	48 (32.0)	36 (24.0)	39 (26.0)	22 (14.7)	5 (3.3)	0.033
	DVS SR 100 mg	147	42 (28.6)	39 (26.5)	38 (25.9)	27 (18.4)	1 (0.7)	0.038
Final	Placebo	150	33 (22.0)	37 (24.7)	41 (27.3)	37 (24.7)	2 (1.3)	--
	DVS SR 50 mg	150	48 (32.0)	35 (23.3)	39 (26.0)	23 (15.3)	5 (3.3)	0.079
	DVS SR 100 mg	147	45 (30.6)	37 (25.2)	36 (24.5)	27 (18.4)	2 (1.4)	0.057

Source: Sponsor's Table 9-3 of CSR

3.1.2.4 Sponsor's Efficacy Results for Other Secondary Endpoints

The sponsor's analysis results based on the LOCF data of the ITT population for the other secondary efficacy variable at the final on-therapy evaluation are shown in Table 3.1.2.5. Among these evaluations, the DVS SR 50-mg dose was statistically significantly better than placebo for the MADRS total score and remission rate at 0.05 significance level

while the DVS SR 100-mg dose did not show any statistically significantly better results than placebo in any four of these measures.

Table 3.1.2.5 Sponsor's Analysis Results (LOCF Data) for Other Secondary Endpoints for Study 332

Endpoints	Placebo (n=150)	DVS SR 50 mg (n=150)	DVS SR 100 mg (n=147)	p-value (vs. Placebo)	
				50-mg	100 mg
Change from Baseline to Final On-Therapy in MADRS Total Score ¹	-12.3 (0.85)	-15.0 (0.85)	-14.3 (0.87)	0.022	0.095
Change from Baseline to Final On-Therapy in CGI-S Total Score ¹	-1.23 (0.10)	-1.47 (0.10)	-1.40 (0.10)	0.074	0.208
HAM-D ₁₇ Response Rate ²	66 (44.0)	80 (54.0)	75 (51.0)	0.098	0.193
HAM-D ₁₇ Remission Rate ²	36 (24.0)	51 (34.0)	46 (31.3)	0.027	0.090

¹ Reported values are least square means and standard errors.

² Reported values are proportions of responders or remitters (i.e., n and %), p-values were based on logistic regression analysis.

3.1.2.5 Statistical Reviewer's Findings and Comments

1. The efficacy evaluation of this study was primarily based on patients' final-on-therapy data. According to the study flow chart in the clinical study report, for subjects who discontinued from the study, safety and efficacy determinations scheduled for day 56 were obtained on the last day the subject took a full dose of study drug or as soon as possible thereafter. Using the data post dosing into the analysis is not suitable and may underestimate the drug's efficacy, although the sponsor's primary efficacy analysis results showed that DVS SR 50 mg is statistically significantly effective in comparison with placebo either by LOCF analysis or the OC analysis. The sponsor has been asked to clarify this point and also perform the MMRM analysis with the unstructured covariance matrix for the cleaned data (See Comment #2 for the details).
2. In addition to the OC analysis, the sponsor also performed some other sensitivity analyses for the primary endpoint (change from baseline to the final on-therapy evaluation of HAM-D₁₇ total score) and reported the p-values in the section of summary of clinical efficacy. One important sensitivity analysis is the MMRM and the sponsor's p-values for the comparisons between DVS SR 50 mg and placebo, and DVS SR 100 mg and placebo are <0.001 and 0.006, respectively. After evaluation, this reviewer found that the sponsor's MMRM analysis results were based on the AR(1) covariance structure, not the unstructured covariance (i.e., UN) that we normally suggest to use. This reviewer found that the reason why the sponsor did not use the UN is due to the convergence problem and this convergence problem was caused by having severely unbalanced data. This reviewer found that only a small portion of patients had Weeks 5 and 7 measurements and the collection of these measurements was not prospectively planned. When patients' Weeks 5 and 7 data were removed, this reviewer found that the MMRM analysis with the unstructured covariance showed p-values, 0.003 and 0.03 for DVS SR 50 mg and DVS SR 100 mg in comparison with placebo, respectively. Table 3.1.2.6 shows the reviewer's MMRM analysis results. On December 21, FDA received the sponsor's response regarding unexpected Weeks 5 and 7 data and the MMRM re-analysis results (SN37). The sponsor explained that the Weeks 5 and 7 data were due to

either patients' late-scheduled-appointments or unscheduled visits. The sponsor's MMRM re-analysis results for the cleaned data were very close to the reviewer's analysis results. After evaluation, this reviewer still thinks that the primary analysis results should be based on the data without the unexpected Weeks 5 and 7 data.

Table 3.1.2.6 Reviewer's MMRM Analysis Results for the Primary Endpoint for Study 332

Therapy Group	n	Adjusted Change from Baseline (SE)	Difference in Adjusted Mean (vs. Placebo)	Overall P-Value	P-Value (vs. Placebo)
Placebo	115	-9.81 (0.62)		0.02	
DVS SR 50 mg	104	-12.46 (0.64)	-2.65		0.003
DVS SR 100 mg	102	-11.73 (0.64)	-1.92		0.03

3. As mentioned in the above comment, this reviewer found that the sponsor's efficacy analysis results were based on data including some unplanned Weeks 5 and 7 measurements for some patients (only ~9% had Week 5 data and ~7% had Week 7 data). Since these patients were not randomly selected, it may bias the analysis results. In addition to the above MMRM analysis this reviewer also performed the primary analysis for the primary endpoint after removing patients' Weeks 5 and/or 7 data and showed the results in Table 3.1.2.7. Note that the p-value for the overall comparisons between three dose groups (DVS SR 50-mg, DVS SR 100mg and placebo) becomes 0.054 and the comparison between DVS SR 100-mg and placebo becomes 0.091 (≥ 0.065 , the sponsor's p-value). According to the pre-specified closed testing procedure, the study becomes negative. The OC analysis results are not affected by this data removal.

Table 3.1.2.7 Reviewer's Analysis Results for Primary Endpoint After Data Removal for Study 332

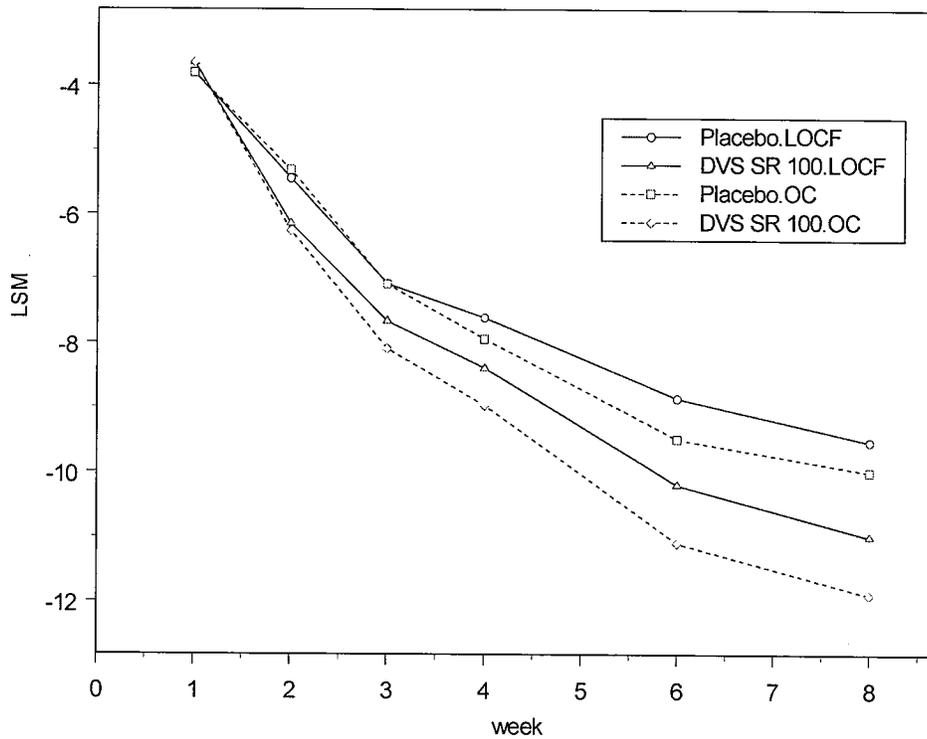
Therapy Group	n	Adjusted Change from Baseline (SE)	Difference in Adjusted Mean (vs. Placebo)	Overall P-Value	P-Value (vs. Placebo)
Placebo	150	-9.57 (0.58)		0.054	
DVS SR 50 mg	150	-11.49 (0.58)	-1.91		0.02
DVS SR 100 mg	147	-10.96 (0.59)	-1.39		0.09

4. For the comparison between DVS SR 100 mg and placebo, since the MMRM and OC analysis results showed completely opposite findings to the LOCF primary analysis results in terms of statistical significance, a reasonable question to ask is whether the conclusion based on the LOCF analysis results should be re-considered.

This reviewer drew a visit-wise plot for both LOCF and OC analysis results on the least square means (Figure 1) and found that the trends of decreasing between the LOCF and OC curves seemed to be similar, although according to Table 3.1.2.1, the major withdrawal reasons seemed to be very different; patients who withdrew from the placebo group were more due to lack of efficacy than in the DVS SR 100 mg group and also those who withdrew from DVS SR 100 mg were more due to adverse event than in Placebo except the unknown reasons.

As we can observe from the figure, it appears that when DVS SR 100mg patients withdrew, their HAMD_17 total scores of change from baseline before their dropout time point were indeed much smaller than those who stayed. It caused big differences between the LOCF and OC curves for patients in DVS SR 100 mg group. That is, similar to the placebo group whose withdrawal reasons were mainly due to lack of efficacy, patients in the drug group who withdrew also showed worse performance. In other words, patients who stayed were those who actually improved. This lack of efficacy in both groups should be taken into consideration for evaluating the drug's efficacy. In this situation, the OC analysis results could be biased. Nevertheless, the overall p-value for the OC analysis is 0.049, which is also close to 0.05. In conclusion, the LOCF analysis results do not seem to be unacceptable in this case.

Figure 1. Least Square Means of ANCOVA for Change from Baseline to Each Visit HAMD₁₇ Total scores for Study 332



3.1.3 Efficacy Results for Study 333

This was a multi-center trial conducted in Europe and South Africa, where there were 44 centers.

3.1.3.1 Patient Dispositions and Baseline Demographic Characteristics

A total of 565 subjects were screened for participation in this study; 80 were screen failures, and 485 subjects were randomly assigned to treatment: 161 were assigned to receive placebo, 166 were assigned to receive DVS SR 50 mg, and 158 were assigned to receive DVS SR 100 mg. Two subjects were excluded from the ITT population because they did not have a primary efficacy evaluation (HAM-D₁₇) while on therapy. So, the ITT population included 483 subjects.

Table 3.1.3.1 summarizes the number of subjects who discontinued treatment by the reasons for withdrawal in each treatment group. Overall, 50 subjects withdrew from the study during the on-therapy period: 13 subjects who received placebo, 17 subjects who received DVS SR 50 mg, and 20 subjects who received DVS SR 100 mg.

Table 3.1.3.1 Number (%) of Subjects Who Discontinued During the On-Therapy Period by Reason for Withdrawal for Study 333

Reason	Placebo n=161	DVS SR 50 mg n=166	DVS SR 100 mg n=158	Total n=485
Total ^a	13 (8.1)	17 (10.2)	20 (12.7)	50 (10.3)
Adverse event	5 (3.1)	8 (4.8)	11 (7.0)	24 (4.9)
Failed to return	0	0	2 (1.3)	2 (0.4)
Investigator request	0	0	2 (1.3)	2 (0.4)
Other	1 (0.6)	0	0	1 (0.2)
Protocol violation	1 (0.6)	4 (2.4)	0	5 (1.0)
Subject request unrelated to study	1 (0.6)	3 (1.8)	4 (2.5)	8 (1.6)
Unsatisfactory response—efficacy	5 (3.1)	2 (1.2)	1 (0.6)	8 (1.6)

Source: Sponsor's Table 8-3 of CSR.

The ITT patients' demographic and baseline characteristics are shown in Table 3.1.3.2. As shown in the table, there was no significant difference between subjects in any characteristic.

Table 3.1.3.2 Demographic and Baseline Characteristics for ITT Population for Study 333

Characteristic	Placebo (n=161)	DVS SR 50 mg (n=164)	DVS SR 100 mg (n=158)
Age (years), n	161	164	158
Mean	45.62	43.96	45.70
Standard deviation	11.55	13.53	12.59
Minimum	19.00	18.00	19.00
Maximum	75.00	78.00	77.00
Age group (years), n (%)			
18-29	18 (11.18)	30 (18.29)	18 (11.39)
30-49	83 (51.55)	77 (46.95)	80 (50.63)
50-64	52 (32.30)	47 (28.66)	48 (30.38)
≥65	8 (4.97)	10 (6.10)	12 (7.59)
Sex, n (%)			
Female	109 (67.70)	115 (70.12)	112 (70.89)
Male	52 (32.30)	49 (29.88)	46 (29.11)
Race, n (%)			
Black or African American	1 (0.62)	0	0
Other: mixed	2 (1.24)	1 (0.61)	2 (1.27)
White	158 (98.14)	163 (99.39)	156 (98.73)
Ethnic origin, n (%)			
Hispanic or Latino	2 (1.24)	2 (1.22)	1 (0.63)
Non-Hispanic	0	0	1 (0.63)
Non-Hispanic and non-Latino	159 (98.76)	162 (98.78)	156 (98.73)

Baseline HAM-D ₁₇ total score			
Mean	24.29	24.26	24.44
Standard deviation	2.60	2.43	2.72
Minimum	20.00	20.00	20.00
Maximum	33.00	32.00	35.00
Median	24.00	24.00	24.00
Baseline CGI-S score			
Mean	4.73	4.66	4.77
Standard deviation	0.68	0.64	0.65
Minimum	4.00	4.00	4.00
Maximum	6.00	6.00	6.00
Median	5.00	5.00	5.00
Baseline CGI-S, n(%)			
(4) Moderately ill	64 (39.8)	70 (42.7)	56 (35.4)
(5) Markedly ill	76 (47.2)	79 (48.2)	83 (52.5)
(6) Severely ill	21 (13.0)	15 (9.1)	19 (12.0)

Source: Sponsor's Table 8-8 of CSR.

3.1.3.2 Sponsor's Efficacy Results for Primary Endpoint

Table 3.1.3.3 shows the sponsor's results for the primary efficacy endpoint, the adjusted mean change from baseline in the HAM-D₁₇ total score at the final on-therapy evaluation. At the final on-therapy evaluation, the adjusted mean change from baseline in the HAM-D₁₇ total score was statistically significantly greater for subjects in the DVS SR 50-mg (p=0.002) and 100-mg (p<0.001) treatment groups compared with the placebo group. At the final on-therapy evaluation, the adjusted mean change from baseline in the HAM-D₁₇ total score was -13.2 in the DVS SR 50-mg group and -13.7 in the DVS SR 100-mg group compared with -10.7 in the placebo group.

Table 3.1.3.3 Sponsor's Analysis Results for HAM-D₁₇ Total Score (by ANCOVA) for Study 333

Analysis	Week of Therapy	Therapy Group	n	Raw Mean Score	Adjusted Change From Baseline	SE	Adjusted Mean (95% CI)	Difference in Adjusted Mean (95% CI) Vs Placebo.	Overall p-Value	p-Value Vs	
										DVS SR 50 mg	DVS SR 100 mg
LOCF	Baseline	Placebo	161	24.3			24.3 (24.3, 24.3)				
		DVS SR 50 mg	164	24.3			24.3 (24.3, 24.3)				
		DVS SR 100 mg	158	24.4			24.3 (24.3, 24.3)				
	Week 1	Placebo	156	21.6	-2.64	0.29	21.7 (21.1, 22.3)		0.788	0.977	0.542
		DVS SR 50 mg	162	21.6	-2.63	0.29	21.7 (21.1, 22.3)	-0.0 (-0.8, 0.8)			0.557
		DVS SR 100 mg	156	22.0	-2.40	0.29	21.9 (21.4, 22.5)	-0.2 (-1.0, 0.5)			
	Week 2	Placebo	161	18.8	-5.53	0.42	18.8 (18.0, 19.6)		0.629	0.386	0.942
		DVS SR 50 mg	164	18.2	-6.03	0.42	18.3 (17.5, 19.1)	0.5 (-0.6, 1.6)			0.429
		DVS SR 100 mg	158	18.8	-5.57	0.42	18.3 (17.9, 19.6)	0.0 (-1.1, 1.2)			
	Week 3	Placebo	161	16.3	-7.93	0.47	16.4 (15.5, 17.3)		0.252	0.168	0.136
		DVS SR 50 mg	164	15.4	-8.81	0.47	15.5 (14.6, 16.4)	0.9 (-0.4, 2.1)			0.898
		DVS SR 100 mg	158	15.4	-8.89	0.48	15.4 (14.5, 16.4)	1.0 (-0.3, 2.2)			
Week 4	Placebo	161	14.9	-9.17	0.52	15.2 (14.1, 16.2)		0.043	0.103	0.014	
	DVS SR 50 mg	164	13.8	-10.3	0.51	14.0 (13.0, 15.0)	1.1 (-0.2, 2.5)			0.389	
	DVS SR 100 mg	158	13.3	-10.9	0.52	13.4 (12.4, 14.4)	1.7 (0.4, 3.1)				
Week 6	Placebo	161	14.1	-10.0	0.57	14.3 (13.2, 15.4)		0.002	0.002	0.002	
	DVS SR 50 mg	164	11.7	-12.4	0.56	11.9 (10.8, 13.0)	2.3 (0.8, 3.8)			0.937	
	DVS SR 100 mg	158	11.8	-12.4	0.57	11.9 (10.8, 13.0)	2.4 (0.9, 3.9)				
Week 8	Placebo	161	13.3	-10.7	0.60	13.7 (12.5, 14.8)		<0.001	0.005	<0.001	
	DVS SR 50 mg	164	10.9	-13.1	0.60	11.2 (10.0, 12.4)	2.4 (0.9, 4.0)			0.511	
	DVS SR 100 mg	158	10.5	-13.7	0.61	10.7 (9.5, 11.9)	3.0 (1.4, 4.6)				
Final on-therapy	Final	Placebo	161	13.3	-10.7	0.61	13.7 (12.5, 14.8)		<0.001	0.002	<0.001
		DVS SR 50 mg	164	10.9	-13.2	0.60	11.2 (10.0, 12.3)	2.5 (0.9, 4.1)			0.498
		DVS SR 100 mg	158	10.4	-13.7	0.61	10.6 (9.4, 11.8)	3.0 (1.4, 4.7)			

Analysis	Week of Therapy	Therapy Group	Raw Mean Score	Adjusted Change From Baseline		Adjusted Mean (95% CI)	Difference in Adjusted Mean (95% CI) Vs Placebo	Overall p-Value	p-Value Vs DVS SR 50 mg	p-Value Vs DVS SR 100 mg	
				n	SE						
Observed	Baseline	Placebo	161	24.3		24.3 (24.3, 24.3)					
		DVS SR 50 mg	164	24.3		24.3 (24.3, 24.3)					
		DVS SR 100 mg	158	24.4		24.3 (24.3, 24.3)					
	Week 1	Placebo	156	21.6	-2.64	0.29	21.7 (21.1, 22.3)		0.788	0.977	0.542
		DVS SR 50 mg	162	21.6	-2.63	0.29	21.7 (21.1, 22.3)	-0.0 (-0.8, 0.8)			0.557
		DVS SR 100 mg	156	22.0	-2.40	0.29	21.9 (21.4, 22.5)	-0.2 (-1.0, 0.5)			
	Week 2	Placebo	155	18.6	-5.75	0.43	18.6 (17.7, 19.4)		0.571	0.291	0.648
		DVS SR 50 mg	155	18.0	-6.36	0.43	18.0 (17.1, 18.8)	0.6 (-0.5, 1.7)			0.562
		DVS SR 100 mg	144	18.5	-6.02	0.44	18.3 (17.4, 19.2)	0.3 (-0.9, 1.4)			
	Week 3	Placebo	154	16.3	-8.09	0.48	16.3 (15.4, 17.3)		0.091	0.042	0.090
		DVS SR 50 mg	148	15.0	-9.42	0.49	15.0 (14.0, 15.9)	1.3 (0.1, 2.6)			0.729
		DVS SR 100 mg	149	15.3	-9.19	0.48	15.2 (14.3, 16.2)	1.1 (-0.2, 2.4)			
	Week 4	Placebo	151	14.8	-9.61	0.51	14.8 (13.8, 15.8)		0.004	0.006	0.004
		DVS SR 50 mg	147	12.8	-11.5	0.52	12.9 (11.9, 13.9)	1.9 (0.6, 3.3)			0.851
		DVS SR 100 mg	140	12.9	-11.6	0.53	12.8 (11.7, 13.8)	2.0 (0.7, 3.4)			
	Week 6	Placebo	137	13.7	-10.9	0.57	13.5 (12.4, 14.6)		<0.001	<0.001	<0.001
		DVS SR 50 mg	144	10.6	-13.9	0.56	10.5 (9.4, 11.6)	3.0 (1.5, 4.4)			0.895
		DVS SR 100 mg	133	11.0	-13.8	0.58	10.6 (9.5, 11.8)	2.9 (1.4, 4.4)			
	Week 8	Placebo	138	12.8	-11.6	0.61	12.8 (11.6, 14.0)		<0.001	<0.001	<0.001
		DVS SR 50 mg	145	9.7	-14.7	0.59	9.7 (8.6, 10.9)	3.1 (1.5, 4.6)			0.533
		DVS SR 100 mg	126	9.4	-15.2	0.63	9.2 (8.0, 10.5)	3.6 (2.0, 5.2)			
	Week >8	Placebo	10	10.3							
		DVS SR 50 mg	5	13.0							
		DVS SR 100 mg	5	7.4							

ANCOVA=analysis of covariance; HAM-D₁₇=Hamilton Rating Scale for Depression, 17-item; ITT=intent-to-treat; LOCF=last-observation-carried-forward; SE=standard error.

Source: Sponsor's Table 9-1 of CSR.

3.1.3.3 Sponsor's Efficacy Results for Key Secondary Endpoint

Table 3.1.3.4 shows the sponsor's results for the key secondary efficacy endpoint, the percentage of subjects with each CGI-I score at the final on-therapy evaluation. At the final on-therapy evaluation, the CGI-I scores differed statistically significantly for subject in the DVS SR 50-mg (p=0.002) and 100-mg (p<0.001) treatment groups compared with the placebo group. At the final on-therapy evaluation, the percentage of subjects with CGI-I scores of 1 (very much improved) or 2 (much improved) was 73% in the DVS SR 50-mg and also DVS SR 100-mg groups compared with 53% in the placebo group.

Table 3.1.3.4 Sponsor's Analysis Results for Secondary Endpoint CGI-I based on the LOCF data for Study 333

Week of Therapy	Therapy Group	n	1 n (%)	2 n (%)	3 n (%)	4 n (%)	5 n (%)	6 n (%)	p-value vs. placebo
Week 8	Placebo	161	56 (34.8)	31 (19.3)	40 (24.8)	24 (14.9)	5 (3.1)	5 (3.1)	-
	DVS SR 50 mg	164	78 (47.6)	41 (25.9)	22 (13.4)	14 (8.5)	6 (3.7)	3 (1.8)	0.003
	DVS SR 100 mg	158	82 (51.9)	34 (21.5)	22 (13.9)	17 (10.8)	3 (1.9)		<0.001
Final	Placebo	161	56 (34.8)	30 (18.6)	40 (24.8)	25 (15.5)	5 (3.1)	5 (3.1)	-
	DVS SR 50 mg	164	79 (48.2)	40 (24.4)	21 (12.8)	15 (9.1)	6 (3.7)	3 (1.8)	0.002
	DVS SR 100 mg	158	82 (51.9)	34 (21.5)	22 (13.9)	17 (10.8)	3 (1.9)		<0.001

Source: Sponsor's Table 9-3 of CSR

3.1.3.4 Sponsor's Efficacy Results for Other Secondary Endpoints

The sponsor's analysis results for other secondary endpoints, change from baseline for MADRS Total score and CGI-S core, response rate based on the HAM-D₁₇ score, and remission defined as HAM-D₁₇ score of 7 or less are shown in Table 3.1.3.5. As shown in the table, at final on-therapy evaluation, both DVS SR 50-mg and 100-mg showed better and statistically significant results in comparison with placebo in the adjusted mean change from baseline in the MADRS total score and also in the CGI-S and also HAM-D₁₇ response rate while for HAM-D₁₇ remission rate, only DVS SR 100-mg showed better and significant results in comparison with placebo at 0.05 significance level.

Table 3.1.3.5 Sponsor's Analysis Results (LOCF Data) for Other Secondary Endpoints for Study 333

Endpoints	Placebo (n=161)	DVS SR 50 mg (n=164)	DVS SR 100 mg (n=157)	p-value	
				50-mg vs. Placebo	100 mg vs. Placebo
Change from Baseline to Final On-Therapy in MADRS Total Score ¹	-13.3 (0.79)	-16.4 (0.78)	-17.5 (0.79)	0.004	<0.001
Change from Baseline to Final On-Therapy in CGI-S Total Score ¹	-1.64 (0.11)	-2.09 (0.11)	-2.19 (0.11)	0.003	<0.001
HAM-D ₁₇ Response Rate ²	80 (49.7)	107 (65.2)	100 (63.3)	0.004	0.011
HAM-D ₁₇ Remission Rate ²	46 (28.6)	61 (37.2)	71 (44.9)	0.099	0.002

¹ Reported values are least square means and standard errors.

² Reported values are proportions of responders or remitters (i.e., n and %) and p-values were based on the logistic regression analysis.

3.1.3.5 Statistical Reviewer's Findings and Comments

1. Comment #1 of Section 3.1.2.5 also applies to Study 333.
2. Similar to Study 332, only some patients had Weeks 5 and 7 data for the primary endpoint although the efficacy conclusion stays the same after removing these patients' Weeks 5 and 7 data. The statistical reviewer's re-analysis results are shown in Table 3.1.3.6. According to the results, both DVS SR 50 mg and DVS SR 100 mg showed statistically significant efficacy findings in comparison with placebo.

Table 3.1.3.6 Reviewer's Analysis Results for Primary Endpoint After Data Removal for Study 333

Therapy Group	n	Adjusted Change from Baseline (SE)	Difference in Adjusted Mean (vs. Placebo)	Overall P-Value	P-Value (vs. Placebo)
Placebo	161	-10.78 (0.60)		<0.001	
DVS SR 50 mg	164	-13.16 (0.59)	-2.37		0.003
DVS SR 100 mg	158	-13.65 (0.60)	-2.86		<0.001

3.2 EVALUATION OF SAFETY

The evaluation of safety was not performed in this review. Please see the clinical review for this evaluation.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 GENDER, RACE AND AGE

The submission did not include the subgroup analysis results only for Studies 332 and 333. In this section, the statistical reviewer's exploratory analysis results are reported for both studies respectively. The adjusted means were obtained from the ANCOVA model with baseline as a covariate, therapy as a factor only. Tables 4.1.1.1 to 4.1.1.3 show the exploratory subgroup analysis results for Study 332 and Tables 4.1.2.1 to 4.1.2.3 are for Study 333.

4.1.1 For Study 332

Table 4.1.1.1 Reviewer's Analysis Results for Age Subgroup Analysis
For Study 332

Therapy Group	n	Adjusted Change from Baseline (SE)	Difference in Adjusted Mean (minus Placebo)
Age < 65			
Placebo	144	-9.41 (0.59)	
DVS SR 50 mg	134	-11.17 (0.61)	-1.76
DVS SR 100 mg	140	-10.85 (0.60)	-1.44
Age ≥ 65			
Placebo	6	-10.14 (3.21)	
DVS SR 50 mg	16	-13.48 (1.92)	-3.34
DVS SR 100 mg	7	-9.93 (2.98)	0.21

Table 4.1.1.2 Reviewer's Analysis Results for Gender Subgroup Analysis
For Study 332

Therapy Group	n	Adjusted Change from Baseline (SE)	Difference in Adjusted Mean (minus Placebo)
Male			
Placebo	55	-9.74 (0.92)	
DVS SR 50 mg	57	-12.20 (0.91)	-2.46
DVS SR 100 mg	69	-11.36 (0.82)	-1.62
Female			
Placebo	95	-9.22 (0.74)	
DVS SR 50 mg	93	-10.94 (0.75)	-1.72
DVS SR 100 mg	78	-10.36 (0.82)	-1.14

Table 4.1.1.3 Reviewer's Analysis Results for Race Subgroup Analysis
For Study 332

Therapy Group	n	Adjusted Change from Baseline (SE)	Difference in Adjusted Mean (minus Placebo)
Race='White'			
Placebo	106	-9.37 (0.69)	
DVS SR 50 mg	110	-11.40 (0.67)	-2.03
DVS SR 100 mg	102	-10.87 (0.70)	-1.50
Race='Black or African American'			
Placebo	29	-9.82 (1.38)	
DVS SR 50 mg	26	-9.21 (1.44)	0.61
DVS SR 100 mg	31	-10.28 (1.33)	-0.46
Race='Others'			
Placebo	15	-9.23 (1.70)	
DVS SR 50 mg	14	-15.61 (1.76)	-6.37
DVS SR 100 mg	14	-11.57 (1.76)	-2.34

4.1.2 For Study 333

Table 4.1.2.1 Reviewer's Analysis Results for Age Subgroup Analysis
For Study 333

Therapy Group	n	Adjusted Change from Baseline (SE)	Difference in Adjusted Mean (minus Placebo)
Age < 65			
Placebo	153	-10.95 (0.58)	
DVS SR 50 mg	154	-13.62 (0.58)	-2.67
DVS SR 100 mg	146	-13.70 (0.60)	-2.76
Age ≥ 65			
Placebo	8	-13.83 (3.06)	
DVS SR 50 mg	10	-10.78 (2.75)	3.05
DVS SR 100 mg	12	-16.46 (2.52)	-2.63

Table 4.1.2.2 Reviewer's Analysis Results for Gender Subgroup Analysis
For Study 333

Therapy Group	n	Adjusted Change from Baseline (SE)	Difference in Adjusted Mean (minus Placebo)
Male			
Placebo	52	-10.46 (0.98)	
DVS SR 50 mg	49	-14.31 (1.01)	-3.85
DVS SR 100 mg	46	-13.46 (1.04)	-3.00
Female			
Placebo	109	-11.32 (0.71)	
DVS SR 50 mg	115	-13.17 (0.69)	-1.85
DVS SR 100 mg	112	-14.07 (0.70)	-2.75

Table 4.1.2.3 Reviewer's Analysis Results for Race Subgroup Analysis
For Study 333

Therapy Group	n	Adjusted Change from Baseline (SE)	Difference in Adjusted Mean (minus Placebo)
Race='White'			
Placebo	158	-11.04 (0.58)	
DVS SR 50 mg	163	-13.56 (0.57)	-2.52
DVS SR 100 mg	156	-13.82 (0.58)	-2.78
Race='Black or African American'			
Placebo	1	-21.0 (.)	
DVS SR 50 mg	0	.	.
DVS SR 100 mg	0	.	.
Race='Others'			
Placebo	2	-14.5 (6.06)	
DVS SR 50 mg	1	18.0 (15.65)	32.5
DVS SR 100 mg	2	-23.5 (6.06)	-9

4.2 OTHER SPECIAL/SUBGROUP POPULATIONS

Study 333 was conducted in Europe and South Africa. To assess the region effect, this reviewer perform the exploratory subgroup analysis by these two regions and the results are shown in Table 4.2.1. According to the results, it was noted that in South Africa patients in DVS SR 50 mg group showed numerically worse performance than patients in placebo group. Nevertheless, there were only total 36 patients from that region. We should interpret results with caution.

Table 4.2.1 Reviewer's Analysis Results for Region Subgroup Analysis
For Study 333

Therapy Group	n	Adjusted Change from Baseline (SE)	Difference in Adjusted Mean (minus Placebo)
Europe			
Placebo	150	-11.04 (0.58)	
DVS SR 50 mg	152	-13.75 (0.58)	-2.71
DVS SR 100 mg	145	-13.99 (0.60)	-2.95
South Africa			
Placebo	11	-11.84 (2.64)	
DVS SR 50 mg	12	-9.57 (2.53)	2.27
DVS SR 100 mg	13	-13.07 (2.43)	-1.23

5. SUMMARY AND CONCLUSIONS

5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE

This reviewer found that for both Studies 332 and 333, a small percentage of patients (around 5% to 10%) had the unplanned Weeks 5 and 7 data. Although the percentages are small, there is still a possibility to bias the analysis results. This reviewer thus performed the re-analyses by removing the unexpected data. The summary of this reviewer's analysis results for the final on-therapy evaluation and the MMRM are shown in Table 2.1.1. Note that the OC analysis results are not affected by this data removal.

According to the reviewer's analysis results, Study 332 appears to be negative but Study 333 still positive although Study 332's overall p-value was very close to 0.05. Based on the sponsor pre-specified closed testing method for multiplicity, none of doses in Study 332 should be further tested after knowing that the overall p-value >0.05 . On the contrary, both 50 mg and 100 mg DVS SR showed statistical significant results in comparison with placebo for Study 333.

5.2 CONCLUSIONS AND RECOMMENDATIONS

Among the two short term studies that the sponsor submitted to evaluate the efficacy of DVS SR 50-mg and DVS SR 100-mg, only Study 333 showed strong evidence of efficacy for both doses. Although the sponsor's analysis results showed that Study 332 supported the efficacy of DVS SR 50-mg, this reviewer had concerns about the cleanness of the data in the sense that some patients have unexpected data. When the unexpected data were removed, the statistical significance of the overall study and DVS SR 50-mg in the primary analysis appears to be diminished. In addition, for Study 332, this reviewer's MMRM analysis results showed larger p-values than the sponsor's although p-values were still less than the nominal significance level 0.05 for both dose levels.

In both studies, the effects of two doses were numerically similar regardless of statistical significance. However, results in Studies 332 and 333 were in favor of the 50 mg dose and 100 mg dose, respectively.

Yeh-Fong Chen, Ph.D.
Mathematical Statistician

cc: NDA 21-992
HFD-130/Dr. Laughren
HFD-130/Dr. Zornberg
HFD-130/Dr. Levin
HFD-130/Ms. Gujral
HFD-700/Dr. Nevius
HFD-700/Ms. Patrician
HFD-710/Dr. Mahjoob
HFD-710/Dr. Hung
HFD-710/Dr. Yang
This review consists of 20 pages. MS Word: C:/yfchen/NDA21992/review.doc

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

/s/

Yeh-Fong Chen
1/8/2008 12:49:00 PM
BIOMETRICS

Peiling Yang
1/8/2008 12:51:02 PM
BIOMETRICS

James Hung
1/10/2008 12:48:56 PM
BIOMETRICS



U.S. Department of Health and Human Services
Food and Drug Administration

Center for Drug Evaluation and Research
Office of Pharmacoepidemiology and Statistical Science
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CARCINOGENICITY STUDIES

NDA: 21-992

Drug Name: DVS-233 (desvenlafaxine succinate)

Applicant: Sanofi-Synthelabo

Biometrics Division: Biometrics Division 6

Statistical Reviewer: Moh-Jee Ng, M.S. (HFD-705)

Concurring Reviewers: Karl Lin, Ph.D. (HFD-705)

Medical Division: Division of Psychiatry Products

Pharmacologist: Linda Fossom, Ph.D. (HFD-130)

Regulatory Manager: Renmeet Grewal, Pharm. D. (HFD-130)

Keywords: NDA review, carcinogenicity

Table of Contents

EXECUTIVE SUMMARY	3
INTRODUCTION	3
REVIEWER'S ANALYSES.....	4
ANALYSIS OF THE RAT DATA.....	6
CONCLUSION OF THE RAT STUDY.....	8
ANALYSIS OF THE MICE DATA	9
CONCLUSION OF THE MOUSE STUDY	11

LIST OF TABLES

TABLE 1: OVERALL DESIGNS OF 2-YEAR CARCINOGENICITY STUDY OF DVS-233 IN RATS AND MICE	4
TABLE 2: ACCUMULATIVE SURVIVAL (%) PRESENTED FOR RATS	6
TABLE 3: NUMBERS OF RATS SURVIVED THE TREATMENT AT WEEK 104	6
TABLE 4: RESULTS OF SIGNIFICANT DOSE-TUMOR TREND TESTS FOR FEMALE RATS.....	6
TABLE 5: RESULTS OF TREND TESTS IN COMBINING TUMORS FOR RATS.....	7
TABLE 6: ACCUMULATIVE SURVIVAL (%) PRESENTED FOR MICE.....	9
TABLE 7: NUMBERS OF MICE SURVIVED THE TREATMENT AT WEEK 104	9
TABLE 8: RESULTS OF SIGNIFICANT DOSE-TUMOR TREND TESTS FOR MICE	8
TABLE 9: RESULTS OF TREND TESTS IN COMBINING TUMORS FOR MICE.....	10

Executive Summary

There were two control groups (CD1 & CD2) and three treated groups, namely low dose (LD), medium dose (MD), and high dose (HD) used in both the rat and the mouse studies. For rats, the dose levels for the LD, MD, and HD groups were 30, 100, and 300 mg/kg/day for males, and 50, 150 and 500 mg/kg/day for females, respectively. For mice, the dose levels for the LD, MD, and HD groups were 50, 150, and 500 mg/kg/day for both males and females. Due to early mortality and a decrease in body weight gain, the high-dose was reduced to 300 mg/kg/day beginning at week 46.

The dose-mortality trend only showed statistically significant in male mice using the Cox test ($p=0.047$) and the Kruskal-Wallis test ($p=0.0032$).

The 2-year rat study showed a statistically significant positive dose-tumor trend in incidence of hepatocellular adenoma in liver ($p=0.0144$, cut off point for rare tumors 0.025) in males.

The 2-year mouse study showed statistically significant positive dose-tumor trend in the incidence of bronchiolo alveolar adenoma in lung ($p=0.0037$), and in the incidence histiocytic sarcoma in multisystemic ($p=0.0049$) in males.

Upon requested by the Executive-CAC, the sponsor submitted historical control data of multisystemic histiocytic sarcoma from two carcinogenicity studies conducted at the same laboratory. The submission included only 2 previous studies completed in the sponsor's laboratory in the last 13 years. Tumor incidences in the four control groups (dual controls in each study) ranged from 0% - 5% in males and 6% - 15% in females, respectively. The respective tumor incidence rates for CD1, CD2, LD, MD and HD groups in the current study (DVS-233) are 0%, 0%, 0%, 3% and 5% for males and 14%, 6%, 8%, 11% and 8% for females.

The sponsor argues that the statistically significant trend and the statistically significant Control-High dose pairwise comparison in incidence rate of this tumor type should be considered as not statistically significant because the incidence rates of the treated groups in the concurrent study are within the range of the historical control data. This reviewer feels that the above sponsor's argument is not valid at least for the following two reasons:

- (1) There are not enough historical control data submitted by the sponsor. Furthermore, the historical control data are from the two very old studies.
- (2) The range of the historical control data is a bad criterion for determining if a statistically significant result is a true significance because the range of historical control data is huge due to the big variability of the data.

Introduction

The objective of this review is to evaluate the oncogenic potential of DVS-233 (Desvenlafaxine Succinate) when administered by oral gavage daily to rats and mice for two years. There were two control groups (CD1 and CD2) and three treated groups, namely low dose (LD), medium dose (MD), and high dose (HD). For rats, the 2 control groups received the vehicle consisting of 0.25% polysorbate 80 and 0.5% methylcellulose mixed in purified water, the dose levels for the LD, MD, and HD groups were 30, 100, and 300 mg/kg/day for males, and 50, 150 and 500 mg/kg/day for females, respectively. For mice, the dose levels for the LD, MD, and HD groups were 50, 150, and 500 mg/kg/day respectively. Due to early mortality and a decrease in body weight gain, the high-dose of the mouse study was reduced to 300 mg/kg/day beginning at week 46. There were 60 rats and 65 mice of each sex in each treatment group. The study design is summarized in Table 1.

Table 1: Overall designs of 2-year carcinogenicity study of DVS-233 in rats and mice

Species	Rat		Mice
Strain	CD [®] (SD) IGS BR		CD-1 [®] (ICR)BR
Route of Administration	Oral		Oral
Dose Unit	mg/kg/day		mg/kg/day
oxymorhorn-HCL (mg/kg/day)	Male	Female	
	0 (CD1)	0 (CD1)	0 (CD1)
	0 (CD2)	0 (CD2)	0 (CD2)
	30 (LD)	50 (LD)	50 (LD)
	100 (MD)	150 (MD)	150 (MD)
	300 (HD)	500 (HD)	500/300* (HD)
Number of Animals/sex/dose	60/sex/dose		65/sex/dose
Length of Study	104 weeks		104 weeks

*: High dose dosage lowered after 45 weeks

Reviewer's Analyses

Analyses of survival and neoplastic data were done using the analysis of carcinogenicity online program -WebCarcin (written by Dr. Ted Guo and Feng Zhou) of Division of Biostatistics II. The test for carcinogenic potential is based on the principles outlined in the Food and Drug Administration's Guidance for Industry: Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceutical (May 2001).

Survival. Homogeneity and trend tests were used to examine the dose-related changes in mortality. Differences in survival distributions among the treatment groups were tested by homogeneity test. A positive trend in the proportion of deaths with respect to the dose levels was tested by trend test. Tests for homogeneity and dose-mortality trends were conducted via the Cox test¹ and the Kruskal-Wallis test². Tables A1-A4 include the numbers of animals at risk, the numbers of animals at deaths, the numbers of animals alive, the cumulative percentages of survival, and the cumulative percentages of deaths by treatment and time intervals. The time intervals used were 0-52, 53-78, 79-91, 93-101, or 93-103 weeks, and the terminal-sacrifice. The actual doses were used as weights. Figures 1-4 present the plots of Kaplan-Meier estimates of the survival distributions of the treatment groups. Tables B1-B4 present results of test of dose-mortality trends and of homogeneity of survival distributions.

Neoplastic Data. The purpose of the analysis of neoplastic data is to determine if there is a positive trend in the proportions of a selected tumor type in a selected organ/tissue with respect to the dose levels. The tumors were classified as either fatal or incidental and were analyzed using the death-rate method³, and the prevalence method, respectively. A combined test was utilized to analyze tumors classified as both fatal and incidental. Multiplicity was addressed employing a decision rule proposed in the guidance. Specifically, positive trends in incidence rates of rare and common tumors were tested at the 0.025 and 0.005 level of significance, respectively. Rare and common tumors were defined based on the tumor rate in the control group. If the tumor rate in the control group was less than 1%, the tumor was classified as rare. Otherwise, the tumor was classified as common. In all analyses, male and female data were analyzed separately for each species. Tables C1-C5 present results of the analysis of dose-tumor trends.

Lastly, to further validate results of negative studies, this reviewer evaluated the number of animals at risk in relation to the adequacy of exposure. Per the guidance document, "a 50% survival rate of the 50 initial animals in any treatment group between weeks 80-90 of a two year study may be considered as a sufficient number and adequate exposure". In addition, this reviewer examined the adequacy of the doses to see if they present a reasonable tumor challenge to the animals. This evaluation was conducted

utilizing criteria outlined by Chu, Cueto, and Ward⁴. Under the criteria, a dose may be considered adequate "if there is a detachable loss in weight gain of up to 10% in a dosed group relative to the controls", or "if dosed animals show a slight increased mortality compared to the control," or " if dose animals exhibit clinical signs or severe histopathologic toxic effects attributed to the chemical."

**Appears This Way
On Original**

¹ Cox, DR: "regression Models and Life tables" *Journal of the Royal Statistical Society, Series B*, 34, 187-220, 1972.

² Gehan, EA: "A Generalized Wilcoxon Test for Comparing K Samples Subject to Unequal Patterns of Censorship" *Biometrika*, 52, 203-223, 1965

³ Peto, R, MC Pike, NE Day, RG Gray, PN Lee, S Parish, J Peto, S Richards, and J Wahrendorf: "Guidelines for Simple Sensitive Significance Tests for Carcinogenic Effects in Long-Term Animal Experiments" *In Long-term and Short-term Screening Assays for Carcinogens: A critical Appraisal*, World Health Organization 1980

⁴ Chu C, C Cueto, and JM Ward: "Factors in the evaluation of 200 National Cancer Institute Carcinogen Bioassays" *Journal of Toxicology and Environmental Health*, 8, 251-280.

Analysis of the Rat Data

Analysis of survival data. The dose-mortality trend was not statistically significant using the Cox test and the Kruskal-Wallis test for both males and females (see appendix Tables B1-B2). Table 2 summarizes the accumulative survivals of the study. The respective accumulative survival rates at the end of the treatment for the CD1, CD2, LD, MD, and HD group were 25%, 30%, 37%, 28%, and 35% in the males, and 27%, 30%, 23%, 27% and 25% in the females, respectively. The survival rates in the females, however, were lower than those in the males, especially the low-dosed group in females (only 14 survival at week 104). Figures 1 and 2 (see appendix) present the survival curves as a function of time for males and females.

Table 2: Accumulative Survival (%) presented for Rats

Deinogest (mg/kg/day) Weeks	Male					Female				
	CD1 0	CD2 0	LD 50	MD 150	HD 300	CD 1 0	CD 2 0	LD 50	MD 150	HD 300
0 - 52	90	92	95	85	93	97	92	93	92	93
53-78	67	63	72	67	68	63	65	58	68	73
79-91	48	42	55	58	63	43	47	43	40	42
92-103	25	30	37	28	35	27	30	23	27	25

Table 3: Numbers of Rats Survived the Treatment at Week 104

Sex	CD1	CD2	LD	MD	HD
Male	15	18	22	17	21
Female	16	18	14	16	15

Analysis of neoplastic data: Table 4 lists the results of significant dose-tumor trend test for male and female rats. The statistical significance for the positive trend test was tested at 0.025 significance level for rare tumors. The statistical significance for the pairwise differences was tested at 0.05 for rare tumors. Tables C1-C2 (see appendix) list the incidence rates of tumors with p-values in testing positive dose-tumor trends for dose groups CD1, CD2, LD, MD, and HD. Since female rats decreased body weight was greater than 10% for the high-dose group, Table C3 (see appendix) lists the dose-tumor trends for dose groups CD1, CD2, LD, and MD (excluded the high-dose group).

Table 4: Results of Significant Dose-Tumor Trend Tests for Rats

DVS-233 / Dose group Organ/ Tumor	CD1	CD2	LD	MD	HD	P-values [†]
Male						
Liver/Hepatocellular adenoma	0	0	0	0	2 ^a	0.0376 ^{†**} 0.0144 [*]

Source data: dataset received on 7/30/2006, analysis data R1M21922 & R1F21992

[†]: p-value of trend tests for dose groups CD1, CD2, LD, MD and HD trend.

^a: p-value for pairwise comparison between the high-dose group and combined control groups.

[‡]: p-value for pairwise comparison between the controls 1 and 2 vs. each individual treatment dose group.

^{*}: statistical significance at 0.025 level. ^{**}: statistical significance at 0.05 level.

^a: one tumor occurred at 70-91 week and the other tumor occurred at the terminal-sacrifice interval

Male rats showed increases in the incidence of hepatocellular adenoma in liver in trend ($p=0.0144$, cut off of 0.025), and in pairwise comparison between the combined controls and the group at 300 mg/kg/day ($p=0.0376$, cut off of 0.05) in males.

Table 5 provides an additional statistical analysis in combining different types of tumors requested by the reviewing pharmacologist. No statistically significant result was found in any of the tumor combinations in both males and females.

Table 5: Results of Trend Tests in Combining Tumors for Rats

DVS-233 / Dose group Organ/ Tumor	CD1	CD2	LD	MD	HD	P-values [†]
Male						
Liver/hepatocellular carcinoma & hepatocellular adenoma	1	2	0	2	4	0.0439
Lung/bronchiolo-alveolar carcinoma & bronchiolo-alveolar adenoma	2	0	0	0	1	0.4419
Adrenal medulla/malignant & benign pheochromocytoma	4	9	13	5	5	0.8903
Mammary gland/adenocarcinoma & fibroadenoma	2	2	1	1	2	0.4956
Multisystem/neurofibrosarcoma & neurofibroma	0	0	2	2	1	0.3363
Multisystemic/fibrosarcoma & fibroma	5	9	9	11	5	0.8289
Pancreas/islet cell carcinoma & islet cell adenoma	5	4	5	6	6	0.3308
Skin/squamous cell papilloma & squamous cell carcinoma	5	4	3	3	1	0.9679
Female						
Adrenal medulla/cortical adenoma & cortical carcinoma	0	1	1	0	0	0.8179
Liver/hepatocellular carcinoma & hepatocellular adenoma	2	0	0	0	0	1.0000
Mammary gland/adenocarcinoma & adenoma	16	21	13	12	17	0.5089
Multisystemic/fibrosarcoma & fibroma	2	3	1	1	2	0.5507
Pancreas/islet cell carcinoma & islet cell adenoma	3	0	1	0	0	0.9084
Pituitary/adenoma & carcinoma	45	47	43	46	47	0.2762
Thyroids/follicular cell carcinoma & follicular cell adenoma	6	7	2	2	5	0.5310
Thyroids/c-cell adenoma & c-cell carcinoma	1	0	1	0	2	0.1123

Source data: dataset received on 7/30/2006, analysis data R2M21922 & R2F21922

[†]: p-value of trend tests for dose groups CD1, CD2, LD, MD and HD trend.

Conclusion of the Rat Study

There were two control groups (CD1 & CD2) and three treated groups, namely low dose (LD), medium dose (MD), and high dose (HD). The dose levels for the LD, MD, and HD groups were 30, 100, and 300 mg/kg/day for males, and 50, 150 and 500 mg/kg/day for females, respectively. The dose-mortality trend was not statistically significant using the Cox test and the Kruskal-Wallis test for both males and females. The respective accumulative survival rates at the end of the treatment for the CD1, CD2, LD, MD, and HD group were 25%, 30%, 37%, 28%, and 35% in the males, and 27%, 30%, 23%, 27% and 25% in the females, respectively. The survival rates in the females, however, were lower than in the males, especially the low-dosed group in females (only 14 surviving at week 104).

The 2-year male rat study showed a statistically significant positive dose-tumor trend in the incidence of hepatocellular adenoma in liver ($p=0.0144$, based on cut off point of 0.025), and a significant pairwise comparison between the combined controls and the group at 300 mg/kg/day ($p=0.0376$, cut off point of 0.05) of the same tumor type.

**Appears This Way
On Original**

Analysis of the Mouse Data

Analysis of survival data. The dose-mortality trend was statistically significant in male mice using the Cox test ($p=0.047$) and the Kruskal-Wallis test ($p=0.0032$). However, the trend was not statistically significant in female mice using the Cox test ($p=0.9205$) and the Kruskal-Wallis test ($p=0.5906$) (see appendix Tables B3-B4). Table 6 summarizes the accumulative survivals of the study. The respective accumulative survival rates at the end of the treatment for the CD1, CD2, LD, MD, and HD group were 46%, 39%, 48%, 43%, and 37% in males, 27%, 25%, 32%, 35% and 31% in females. Each group had at least 16 mice surviving to the scheduled sacrifice at week 104 (see Table 7). Figures 3 and 4 (see appendix) present the survival curves as a function of time for males and females.

Table 6: Accumulative Survival (%) presented for Mice

Sex	Male					Female				
	CD1	CD2	LD	MD	HD	CD1	CD2	LD	MD	HD
DVS-233 (mg/kg/day)	0	0	50	150	500/300	0	0	50	150	500/300
Weeks 0 - 52	94	97	97	99	79	99	95	95	95	83
53 - 78	80	80	82	77	57	74	66	71	75	65
79 - 91	60	62	60	59	43	46	45	59	49	46
92 - 103	46	39	48	43	37	27	25	32	35	31

Table 7: Numbers of Mice Survived the Treatment at Week 104

DVS-233 (mg/kg/day) Sex	CD1	CD2	LD	MD	HD
	0	0	50	150	500/300
Male	30	25	31	28	24
Female	17	16	21	23	20

Analysis of neoplastic data: Table 8 lists the results of significant dose-tumor trend test for male and female mice. The statistical significance for the positive trend test was tested at 0.025 and 0.005 significance levels for rare and common tumors, respectively. The statistical significance for the pairwise differences was tested at 0.01 and 0.05 for rare and common tumors, respectively. Tables C4-C5 (see appendix) list the incidence rates of tumors with p-values in testing positive dose-tumor trends.

Table 8: Results of Significant Dose-Tumor Trend Tests for Mice

DVS-233 / Dose group Organ/ Tumor	CD1	CD2	LD	MD	HD	P-values†
	0	0	50	150	500/300	
Male						
Lung/M-bronchiolo alveolar adenoma	4	5	10	10	12	0.0037*
Multisystemic/M-histiocytic sarcoma	0	0	0	2	3	0.0049**
Female						
Ovaries/B-cystadenoma	0	3	7	0	2	0.6773

Source data: dataset received on 7/30/2006, analysis data M1M21922 & M1F21922

†: p-value of trend tests for dose groups CD1, CD2, LD, MD and HD trend.

‡: p-value for pairwise comparison between the controls 1 and 2 vs. each individual treatment dose group.

*: statistical significance at 0.005 level. **: statistical significance at 0.025 level.

: statistical significance at 0.01 level. *: statistical significance at 0.05 level.

Male mice showed increases in the incidence of the following neoplastic findings:

- in the incidence of bronchiolo alveolar adenoma in lung in trend ($p=0.0037$, cut off point of 0.005), and in pairwise comparison between the combined controls and the group at 500/300 mg/kg/day ($p=0.0076$, cut off point of 0.01).
- in the incidence of histiocytic sarcoma in multisystemic in trend ($p=0.0049$, cut off point of 0.025), in pairwise comparison between the combined controls and the group at 500/300 mg/kg/day ($p=0.0217$, cut off point of 0.05).

The positive trend was not statistically significance in the incidence of cystadenoma in ovaries female mice ($p=0.6673$, cut off point of 0.005). However, the pairwise comparison between the combined controls and the group at 50 mg/kg/day is significant ($p=0.0076$, cut off point of 0.01).

Table 9 provides an additional statistical analysis in combining different types of tumors requested by the reviewing pharmacologist. No statistically significant result was found in any of the tumor combinations in both males and females.

Table 9: Results of Trend Tests in Combining Tumors for Mice

DVS-233 / Dose group Organ/ Tumor	CD1 0	CD2 0	LD 50	MD 150	HD 500/300	P-values [†]
Male						
Adrenal cortex/cortical carcinoma & cortical adenoma	4	1	1	2	0	0.8902
Kidneys/tubular carcinoma & tubular adenoma	2	0	0	0	1	0.5224
Liver/hepatocellular adnoma & hepatocellular carcinoma	22	21	14	20	7	0.9748
Lung/bronchiolo-alveolar adenoma & bronchiolo-alveolar carcinom	15	18	17	18	16	0.2803
Female						
Liver/hepatocellular adnoma & hepatocellular carcinoma	1	1	1	2	0	0.7318
Lung/bronchiolo-alveolar adenoma & bronchiolo-alveolar carcinom	11	11	10	10	8	0.7549
Ovaries/cystadenoma & cystadeno papillary, bilateral	0	3	7	0	4	0.2889
Ovaries/cystadenoma & cystadeno carcinoma & cystadeno papillary, bilateral	0	3	7	0	5	0.1979
Pituitary/adenoma & carcinoma	7	0	2	0	5	0.3225

Source data: dataset received on 7/30/2006, analysis data M2M21992 & M2F21992

[†]: p-value of trend tests for dose groups CD1, CD2, LD, MD and HD trend.

Conclusion of the Mouse Study

There were two control groups (CD1 & CD2) and three treated groups, namely low dose (LD), medium dose (MD), and high dose (HD) used in the mouse study. The dose levels for the LD, MD, and HD groups were 30, 100, and 300 mg/kg/day for males, and 50, 150 and 500 mg/kg/day for females, respectively. Due to early mortality and a decrease in body weight gain, the high-dose was reduced to 300 mg/kg/day beginning at week 46. The dose-mortality trend was statistically significant in male mice using the Cox test ($p=0.047$) and the Kruskal-Wallis test ($p=0.0032$). However, the trend was not statistically significant in female mice using the Cox test ($p=0.9205$) and the Kruskal-Wallis test ($p=0.5906$). The respective accumulative survival rates at the end of the treatment for the CD1, CD2, LD, MD, and HD group were 46%, 39%, 48%, 43%, and 37% in males, 27%, 25%, 32%, 35% and 31% in females. Each group had at least 16 mice surviving to the scheduled sacrifice at week 104.

The 2-year male mouse study showed a statistically significant positive dose-tumor trend in the incidence of bronchiolo alveolar adenoma in lung in trend ($p= 0.0037$, based on cut off point of 0.005), and a significant pairwise comparison in the incidence between the combined controls and the group at 500/300 mg/kg/day ($p=0.0076$, cut off point of 0.01) of the same tumor. Also the trend in the incidence of histiocytic sarcoma in multisystemic ($p = 0.0049$, based on cut off point of 0.025), and the pairwise comparison between the combined controls and the group at 500/300 mg/kg/day ($p =0.0217$, based on cut off point of 0.05) of the same tumor are statistically significant.

In the 2-year female mouse study, the positive dose-tumor trend in the incidence of cystadenoma in ovaries was not statistically significant ($p=0.6673$, based on cut off point of 0.005). However, the pairwise comparison in the incidence between the combined controls and the group at 50 mg/kg/day was statistically significant ($p =0.0076$, based on cut off point of 0.01) of the same tumor.

**Appears This Way
On Original**

Appendices

Table A1: Analysis of Mortality Data for Male Rats by Treatment and Time

Analysis of Mortality	No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality	
Control 1 0 mg/kg/day	0-52	60	6	54	90.0	10.0
	53-78	54	14	40	66.7	33.3
	79-91	40	11	29	48.3	51.7
	92-103	29	14	15	25.0	75.0
	Terminal Sacrifice	15	15	0		
Control 2 0 mg/kg/day	0-52	60	5	55	91.7	8.3
	53-78	55	17	38	63.3	36.7
	79-91	38	13	25	41.7	58.3
	92-103	25	7	18	30.0	70.0
	Terminal Sacrifice	18	18	0		
Low Dose 50 mg/kg/day	0-52	60	3	57	95.0	5.0
	53-78	57	14	43	71.7	28.3
	79-91	43	10	33	55.0	45.0
	92-103	33	11	22	36.7	63.3
	Terminal Sacrifice	22	22	0		
Med Dose 150 mg/kg/day	0-52	60	9	51	85.0	15.0
	53-78	51	11	40	66.7	33.3
	79-91	40	5	35	58.3	41.7
	92-103	35	18	17	28.3	71.7
	Terminal Sacrifice	17	17	0		
High Dose 500mg/kg/day	0-52	60	4	56	93.3	6.7
	53-78	56	15	41	68.3	31.7
	79-91	41	3	38	63.3	36.7
	92-103	38	17	21	35.0	65.0
	Terminal Sacrifice	21	21	0		

Source data: dataset received on 7/30/2006, analysis data R1M21922

Table A2: Analysis of Mortality Data for Female Rats by Treatment and Time

Analysis of Mortality	No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality	
Control 1 0 mg/kg/day	0-52	60	2	58	96.7	3.3
	53-78	58	20	38	63.3	36.7
	79-91	38	12	26	43.3	56.7
	92-101	26	10	16	26.7	73.3
	Terminal Sacrifice	16	16	0		
Control 2 0 mg/kg/day	0-52	60	5	55	91.7	8.3
	53-78	55	16	39	65.0	35.0
	79-91	39	11	28	46.7	53.3
	92-101	28	10	18	30.0	70.0
	Terminal Sacrifice	18	18	0		
Low Dose 30 mg/kg/day	0-52	60	4	56	93.3	6.7
	53-78	56	21	35	58.3	41.7
	79-91	35	9	26	43.3	56.7
	92-101	26	12	14	23.3	76.7
	Terminal Sacrifice	14	14	0		
Med Dose 100 mg/kg/day	0-52	60	5	55	91.7	8.3
	53-78	55	14	41	68.3	31.7
	79-91	41	17	24	40.0	60.0
	92-101	24	8	16	26.7	73.3
	Terminal Sacrifice	16	16	0		
High Dose 300mg/kg/day	0-52	60	4	56	93.3	6.7
	53-78	56	12	44	73.3	26.7
	79-91	44	19	25	41.7	58.3
	92-101	25	10	15	25.0	75.0
	Terminal Sacrifice	15	15	0		

Source data: dataset received on 7/30/2006, analysis data R1F21922

Table A3: Analysis of Mortality Data for Male Mice by Treatment and Time

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
Control 1 0 mg/kg/day	0-52	65	4	61	93.8	6.2
	53-78	61	9	52	80.0	20.0
	79-91	52	13	39	60.0	40.0
	92-103	39	9	30	46.2	53.8
	Terminal Sacrifice	30	30	0		
Control 2 0 mg/kg/day	0-52	65	2	63	96.9	3.1
	53-78	63	11	52	80.0	20.0
	79-91	52	12	40	61.5	38.5
	92-103	40	15	25	38.5	61.5
	Terminal Sacrifice	25	25	0		
Low Dose 50 mg/kg/day	0-52	65	2	63	96.9	3.1
	53-78	63	10	53	81.5	18.5
	79-91	53	14	39	60.0	40.0
	92-103	39	8	31	47.7	52.3
	Terminal Sacrifice	31	31	0		
Med Dose 150 mg/kg/day	0-52	65	1	64	98.5	1.5
	53-78	64	14	50	76.9	23.1
	79-91	50	12	38	58.5	41.5
	92-103	38	10	28	43.1	56.9
	Terminal Sacrifice	28	28	0		
High Dose 500/300mg/kg/day ^a	0-52	65	14	51	78.5	21.5
	53-78	51	14	37	56.9	43.1
	79-91	37	9	28	43.1	56.9
	92-103	28	4	24	36.9	63.1
	Terminal Sacrifice	24	24	0		

a: At week 46, the high dosage was decreased from 500 to 300 mg/kg/day
Source data: dataset received on 7/31/2006, analysis data M1M21992

Table A4: Analysis of Mortality Data for Female Mice by Treatment and Time

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
Control 1 0 mg/kg/day	0-52	65	1	64	98.5	1.5
	53-78	64	16	48	73.8	26.2
	79-91	48	18	30	46.2	53.8
	92-103	30	13	17	26.2	73.8
	Terminal Sacrifice	17	17	0		
Control 2 0 mg/kg/day	0-52	65	3	62	95.4	4.6
	53-78	62	19	43	66.2	33.8
	79-91	43	14	29	44.6	55.4
	92-103	29	13	16	24.6	75.4
	Terminal Sacrifice	16	16	0		
Low Dose 50 mg/kg/day	53-78	65	19	46	70.8	29.2
	79-91	46	8	38	58.5	41.5
	92-103	38	17	21	32.3	67.7
	Terminal Sacrifice	21	21	0		
Med Dose 150 mg/kg/day	0-52	65	3	62	95.4	4.6
	53-78	62	13	49	75.4	24.6
	79-91	49	17	32	49.2	50.8
	92-103	32	9	23	35.4	64.6
	Terminal Sacrifice	23	23	0		
High Dose 500/300mg/kg/day ^a	0-52	65	11	54	83.1	16.9
	53-78	54	12	42	64.6	35.4
	79-91	42	12	30	46.2	53.8
	92-103	30	10	20	30.8	69.2
	Terminal Sacrifice	20	20	0		

a: At week 46, the high dosage was decreased from 500 to 300 mg/kg/day
Source data: dataset received on 4/11/2005, analysis data M1F21922

Figure 1: Kaplan-Meier Survival Curve of the 2-year Oral Carcinogenicity Study of DVS-233 in Male Rats

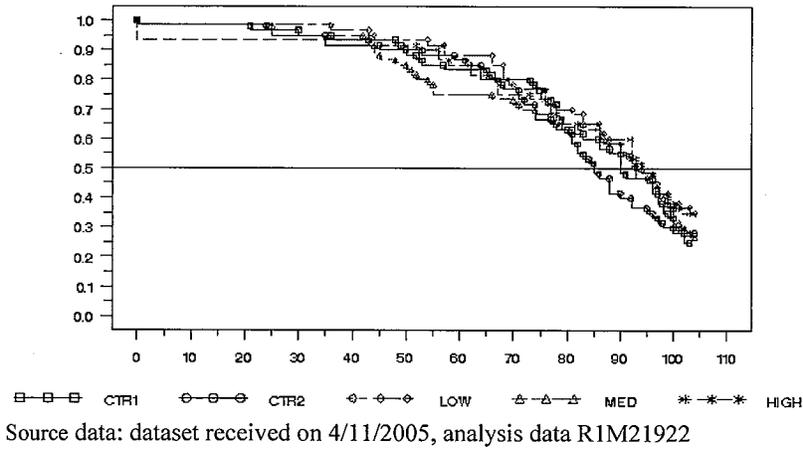


Figure 2: Kaplan-Meier Survival Curve of the 2-year Oral Carcinogenicity Study of DVS-233 in Female Rats

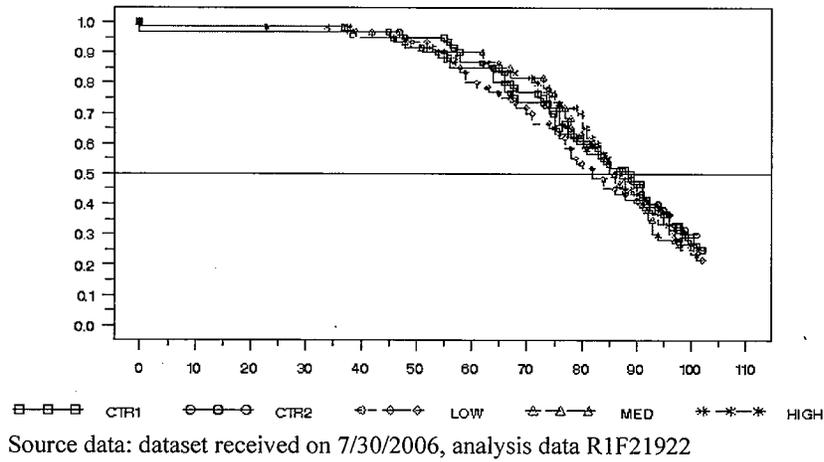


Figure 3: Kaplan-Meier Survival Curve of the 2-year Oral Carcinogenicity Study of DVS-233 in Male Mice

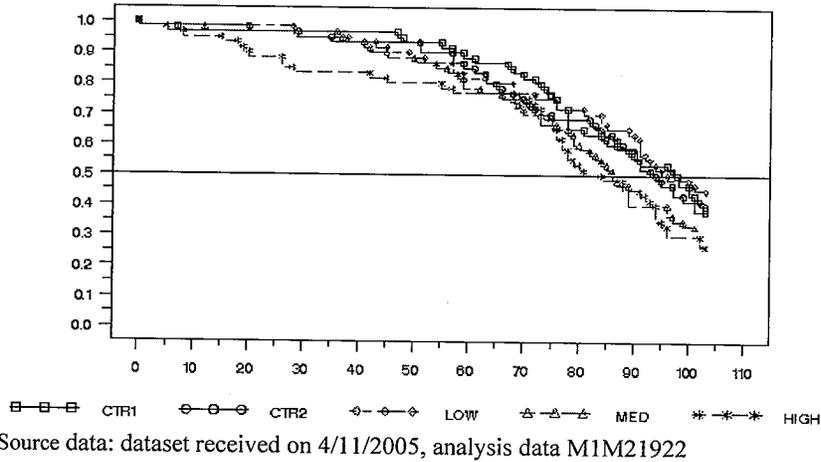


Figure 4: Kaplan-Meier Survival Curve of the 2-year Oral Carcinogenicity Study of DVS-233 in Female Mice

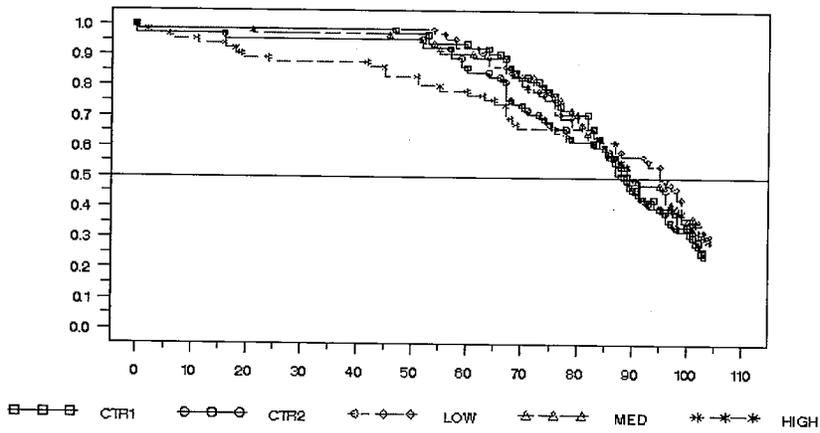


Table B1: Analysis of Dose-Mortality Trend for Male Rats

Time-Adjusted Trend Test	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Dose-Mortality Trend	0.6731	0.4120	0.5576	0.4552
Homogeneity	2.3505	0.6716	2.3563	0.6705

P-value in bold areas showed statistically significant at 0.05 level.

Table B2: Analysis of Dose-Mortality Trend for Female Rats

Time-Adjusted Trend Test	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Dose-Mortality Trend	0.0010	0.9746	0.0807	0.7763
Homogeneity	0.6834	0.9534	0.9446	0.9181

P-value in bold areas showed statistically significant at 0.05 level.

Table B3: Analysis of Dose-Mortality Trend for Male Mice

Time-Adjusted Trend Test	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Dose-Mortality Trend	3.9461	0.0470	8.7053	0.0032
Homogeneity	6.0905	0.1925	11.8390	0.0186

P-value in bold areas showed statistically significant at 0.05 level.

Table B4: Analysis of Dose-Mortality Trend for Female Mice

Time-Adjusted Trend Test	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Dose-Mortality Trend	0.0100	0.9205	0.2894	0.5906
Homogeneity	2.1266	0.7125	2.8116	0.5898

P-value in bold areas showed statistically significant at 0.05 level.

Table C1: Report on Test for Positive Dose-Tumor Trends in Male Rats

Organ Name	Tumor Name	CTR1	CTR2	LOW	MED	HIGH	P-Value
CERVICAL CORD	M-ASTROCYTOMA	0	0	0	1	0	0.4196
ADRENAL CORTEX	B-CORTICAL ADENOMA	0	1	0	1	1	0.3371
HARDERIAN GLAND	B-ADENOMA	0	0	0	1	0	0.4815
HEART	M-SCHWANNOMA, ENDOCARDIAL	0	1	0	0	0	1.0000
KIDNEYS	M-RENAL MESENCHYMAL TUMOR	1	0	0	0	0	1.0000
LIVER	M-HEPATOCELLULAR CARCINOMA	1	2	0	2	2	0.3615
LIVER	B-HEPATOCELLULAR ADENOMA	0	0	0	0	2	0.0144
LUNG	M-BRONCHIOLO-ALVEOLAR CARCINOM	2	0	0	0	0	1.0000
LUNG	B-BRONCHIOLO-ALVEOLAR ADENOMA	0	0	0	0	1	0.2113
ADRENAL MEDULLA	M-MALIGNANT PHEOCHROMOCYTOMA	0	3	1	0	2	0.4894
ADRENAL MEDULLA	B-BENIGN PHEOCHROMOCYTOMA	4	6	12	5	3	0.9180
ADRENAL MEDULLA	B-GANGLIONEUROMA	1	0	0	0	0	1.0000
MAMMARY GLAND	M-ADENOCARCINOMA	0	1	0	0	1	0.3664
MAMMARY GLAND	B-FIBROADENOMA	2	1	1	1	1	0.6703
MULTISYSTEMIC	M-HEMANGIOSARCOMA	1	2	0	0	0	0.8874
MULTISYSTEMIC	M-HISTIOCYTIC SARCOMA	0	1	1	3	1	0.3894
MULTISYSTEMIC	M-NEUROFIBROSARCOMA	0	0	2	1	1	0.3163
MULTISYSTEMIC	M-LYMPHOMA	1	1	0	1	3	0.0429
MULTISYSTEMIC	M-MESOTHELIOMA	1	0	0	1	1	0.2843
MULTISYSTEMIC	M-UNDIFFERENTIATED SARCOMA	1	1	1	1	0	0.8455
MULTISYSTEMIC	B-NEUROFIBROMA	0	0	0	1	0	0.3810
MULTISYSTEMIC	M-LIPOSARCOMA	1	0	0	0	0	1.0000
MULTISYSTEMIC	M-OSTEOSARCOMA	0	0	0	1	0	0.4086
MULTISYSTEMIC	M-FIBROSARCOMA	2	5	3	6	3	0.5799
MULTISYSTEMIC	B-FIBROMA	3	4	6	5	2	0.8642
PANCREAS	M-ISLET CELL CARCINOMA	2	0	3	2	2	0.3906
PANCREAS	B-ISLET CELL ADENOMA	3	4	2	4	4	0.3649
PANCREAS	B-ACINAR CELL ADENOMA	1	1	0	1	0	0.7713
PARATHYROIDS	B-ADENOMA	0	1	3	0	0	0.8902
PITUITARY	B-ADENOMA	29	30	24	27	23	0.9305
SEMINAL VESICLE	M-LEIOMYOSARCOMA	0	0	0	1	0	0.1750
SKIN	B-SQUAMOUS CELL PAPILLOMA	3	3	2	3	1	0.8702
SKIN	M-MYXOSARCOMA	0	0	0	0	1	0.1992
SKIN	M-LEIOMYOSARCOMA	0	0	0	0	1	0.2078
SKIN	M-SQUAMOUS CELL CARCINOMA	2	1	1	0	0	0.9848
SKIN	B-LIPOMA	2	2	2	6	0	0.8810
SKIN	B-TRICHOEPITHELIOMA	0	1	0	0	1	0.3955
SKIN	M-MALIGNANT BASAL CELL TUMOR	1	0	1	0	0	0.8694

SQUAMOUS STOMACH	B-SQUAMOUS CELL PAPILLOMA	2	1	0	0	0	1.0000
STOMACH	M-ADENOCARCINOMA	0	0	0	1	0	0.4022
TESTES	B-INTERSTITIAL CELL ADENOMA	2	2	5	2	5	0.2383
THYMUS	M-MALIGNANT THYMOMA	0	0	0	0	1	0.1692
THYROIDS	B-C-CELL ADENOMA	5	5	8	6	3	0.8561
THYROIDS	M-C-CELL CARCINOMA	1	0	1	1	2	0.1604
THYROIDS	B-FOLLICULAR CELL ADENOMA	0	1	1	1	1	0.3578
TONGUE	M-SQUAMOUS CELL CARCINOMA	1	0	0	0	0	1.0000
URINARY BLADDER	M-CARCINOMA	0	0	0	0	1	0.2192
URINARY BLADDER	B-PAPILLOMA	0	0	0	0	1	0.2258
BRAIN	M-ASTROCYTOMA	2	0	0	2	1	0.4695

Note: The check mark  indicates statistically significant test results, based on the decision rule of FDA .CDER.Divisions of Biometrics.

source data: dataset received on 7/30/2006, analysis data R1M21922

Appears This Way
On Original

Table C2: Report on Test for Positive Dose-Tumor Trends in Female Rats

Organ Name	Tumor Name	CTR1	CTR2	LOW	MED	HIGH	P-Value
CERVICAL CORD	M-ASTROCYTOMA	0	0	1	0	0	0.5978
ADRENAL CORTEX	B-CORTICAL ADENOMA	0	0	1	0	0	0.5696
ADRENAL CORTEX	M-CORTICAL CARCINOMA	0	1	0	0	0	1.0000
LIVER	M-HEPATOCELLULAR CARCINOMA	1	0	0	0	0	1.0000
LIVER	B-HEPATOCELLULAR ADENOMA	1	0	0	0	0	1.0000
LUNG	M-BRONCHIOLO-ALVEOLAR CARCINOM	1	2	0	0	0	1.0000
ADRENAL MEDULLA	B-BENIGN PHEOCHROMOCYTOMA	1	1	0	0	0	1.0000
MAMMARY GLAND	M-ADENOCARCINOMA	14	17	12	10	15	0.4686
MAMMARY GLAND	B-ADENOMA	2	5	2	2	2	0.7097
MAMMARY GLAND	M-ADENOCARCINOMA ARISING IN A	0	2	0	1	0	0.7686
MAMMARY GLAND	B-FIBROADENOMA	31	37	20	22	12	0.9999
MULTISYSTEMIC	M-HISTIOCYTIC SARCOMA	1	3	2	1	1	0.8195
MULTISYSTEMIC	M-LYMPHOMA	0	1	1	0	0	0.8449
MULTISYSTEMIC	M-MESOTHELIOMA	1	0	1	0	0	0.8179
MULTISYSTEMIC	M-UNDIFFERENTIATED SARCOMA	0	0	0	1	0	0.3980
MULTISYSTEMIC	M-FIBROSARCOMA	0	0	1	0	0	0.5872
MULTISYSTEMIC	B-FIBROMA	2	3	0	1	2	0.4905
OVARIES	B-BENIGN GRANULOSA CELL TUMOR	0	0	1	0	1	0.2135
PANCREAS	M-ISLET CELL CARCINOMA	1	0	0	0	0	1.0000
PANCREAS	B-ISLET CELL ADENOMA	2	0	1	0	0	0.9332
PANCREAS	B-ACINAR CELL ADENOMA	2	0	0	0	1	0.4765
PARATHYROIDS	B-ADENOMA	1	1	0	0	0	1.0000
PITUITARY	B-ADENOMA	45	43	41	43	43	0.4591
PITUITARY	M-CARCINOMA	0	4	2	3	4	0.1610
SKELETAL MUSCLE	M-RHABDOMYOSARCOMA	1	0	0	0	0	1.0000
SKIN	B-HIBERNOMA	0	0	0	0	1	0.1875
THYROIDS	B-C-CELL ADENOMA	6	6	1	2	4	0.6175
THYROIDS	M-FOLLICULAR CELL CARCINOMA	0	0	0	0	2	0.0548
THYROIDS	M-C-CELL CARCINOMA	0	1	1	0	2	0.1172
THYROIDS	B-FOLLICULAR CELL ADENOMA	1	0	1	0	0	0.8179
TONGUE	B-SQUAMOUS CELL PAPILLOMA	0	0	1	0	0	0.5610
URINARY BLADDER	B-PAPILLOMA	0	0	0	0	1	0.1899
UTERUS	B-ENDOMETRIAL STROMAL POLYP	2	2	4	3	1	0.7957
UTERUS	M-LEIOMYOSARCOMA	0	1	0	0	0	1.0000
VAGINA	B-POLYP	1	0	0	0	1	0.3497
BRAIN	M-ASTROCYTOMA	0	1	1	0	0	0.8371

Table C3: Report on Test for Positive Dose-Tumor Trends in Female Rats
(excluded the high-dose group)

Organ Name	Tumor Name	CTR1	CTR2	LOW	MED	P-Value (Exact Method)
CERVICAL CORD	M-ASTROCYTOMA	0	0	1	0	0.5000
ADRENAL CORTEX	B-CORTICAL ADENOMA	0	0	1	0	0.4687
ADRENAL CORTEX	M-CORTICAL CARCINOMA	0	1	0	0	1.0000
LIVER	M-HEPATOCELLULAR CARCINOMA	1	0	0	0	1.0000
LIVER	B-HEPATOCELLULAR ADENOMA	1	0	0	0	1.0000
LUNG	M-BRONCHIOLO-ALVEOLAR CARCINOM	1	2	0	0	1.0000
ADRENAL MEDULLA	B-BENIGN PHEOCHROMOCYTOMA	1	1	0	0	1.0000
MAMMARY GLAND	M-ADENOCARCINOMA	14	17	12	10	0.8870
MAMMARY GLAND	B-ADENOMA	2	5	2	2	0.7619
MAMMARY GLAND	M-ADENOCARCINOMA ARISING IN A	0	2	0	1	0.5587
MAMMARY GLAND	B-FIBROADENOMA	31	37	20	22	0.9726
MULTISYSTEMIC	M-HISTIOCYTIC SARCOMA	1	3	2	1	0.7621
MULTISYSTEMIC	M-LYMPHOMA	0	1	1	0	0.7564
MULTISYSTEMIC	M-MESOTHELIOMA	1	0	1	0	0.7217
MULTISYSTEMIC	M-UNDIFFERENTIATED SARCOMA	0	0	0	1	0.2500
MULTISYSTEMIC	M-FIBROSARCOMA	0	0	1	0	0.4767
MULTISYSTEMIC	B-FIBROMA	2	3	0	1	0.8627
OVARIES	B-BENIGN GRANULOSA CELL TUMOR	0	0	1	0	0.4930
PANCREAS	M-ISLET CELL CARCINOMA	1	0	0	0	1.0000
PANCREAS	B-ISLET CELL ADENOMA	2	0	1	0	0.8706
PANCREAS	B-ACINAR CELL ADENOMA	2	0	0	0	1.0000
PARATHYROIDS	B-ADENOMA	1	1	0	0	1.0000
PITUITARY	B-ADENOMA	45	43	41	43	0.3777
PITUITARY	M-CARCINOMA	0	4	2	3	0.3366
SKELETAL MUSCLE	M-RHABDOMYOSARCOMA	1	0	0	0	1.0000
THYROIDS	B-C-CELL ADENOMA	6	6	1	2	0.9595
THYROIDS	M-C-CELL CARCINOMA	0	1	1	0	0.7276
THYROIDS	B-FOLLICULAR CELL ADENOMA	1	0	1	0	0.7217
TONGUE	B-SQUAMOUS CELL PAPILLOMA	0	0	1	0	0.4857
UTERUS	B-ENDOMETRIAL STROMAL POLYP	2	2	4	3	0.2847
UTERUS	M-LEIOMYOSARCOMA	0	1	0	0	1.0000
VAGINA	B-POLYP	1	0	0	0	1.0000
BRAIN	M-ASTROCYTOMA	0	1	1	0	0.7502

Table C4: Report on Test for Positive Dose-Tumor Trends in Male Mice

Organ Name	Tumor Name	CTR1	CTR2	LOW	MED	HIGH	P-Value
ADRENAL CORTEX	M-CORTICAL CARCINOMA	1	0	0	0	0	1.0000
ADRENAL CORTEX	B-CORTICAL ADENOMA	3	1	1	2	0	0.8598
CECUM	M-ADENOCARCINOMA	0	0	0	0	1	0.1646
GALLBLADDER	B-PAPILLOMA	0	0	0	0	1	0.1603
ADRENAL MEDULLA	B-BENIGN PHEOCHROMOCYTOMA	0	0	2	1	1	0.3255
HARDERIAN GLAND	B-ADENOMA	6	7	8	7	4	0.6774
KIDNEYS	M-TUBULAR CARCINOMA	2	0	0	0	0	1.0000
KIDNEYS	B-TUBULAR ADENOMA	0	0	0	0	1	0.1739
LIVER	B-HEPATOCELLULAR ADENOMA	15	13	13	12	4	0.9783
LIVER	M-HEPATOCELLULAR CARCINOMA	7	9	1	9	3	0.7177
LUNG	B-BRONCHIOLO-ALVEOLAR ADENOMA	4	5	10	10	12	0.0037 !
LUNG	M-BRONCHIOLO-ALVEOLAR CARCINOM	12	14	7	8	5	0.9479
MULTISYSTEMIC	M-SCHWANNOMA	0	1	0	1	0	0.5743
MULTISYSTEMIC	M-MESOTHELIOMA	1	0	0	0	0	1.0000
MULTISYSTEMIC	M-HEMANGIOSARCOMA	5	6	3	4	5	0.3081
MULTISYSTEMIC	M-FIBROSARCOMA	0	0	1	0	0	0.5628
MULTISYSTEMIC	M-HISTIOCYTIC SARCOMA	0	0	0	2	3	0.0049 !
MULTISYSTEMIC	M-LYMPHOMA	2	6	6	6	4	0.3266
MULTISYSTEMIC	M-OSTEOSARCOMA	0	0	2	0	0	0.6415
MULTISYSTEMIC	M-UNDIFFERENTIATED SARCOMA	0	1	1	0	0	0.8290
PITUITARY	B-ADENOMA	1	0	0	0	0	1.0000
SEMINAL VESICLE	B-ADENOMA	1	0	0	0	0	1.0000
SKIN	M-CARCINOMA, BASAL CELL	1	0	0	0	0	1.0000
SPLEEN	B-HEMANGIOMA	1	0	0	0	0	1.0000
SQUAMOUS STOMACH	M-SQUAMOUS CELL CARCINOMA	0	0	2	0	0	0.6625
TESTES	B-INTERSTITIAL CELL ADENOMA	2	3	2	3	2	0.4158
TESTES	B-PAPILLARY ADENOMA, RETE TEST	1	0	0	0	0	1.0000
THYROIDS	B-FOLLICULAR ADENOMA	1	0	3	1	1	0.4499
BRAIN	M-ASTROCYTOMA	1	0	0	0	0	1.0000
BRAIN	B-MENINGIOMA	0	0	0	1	0	0.3539
BRAIN	M-MENINGIOMA	1	0	0	0	0	1.0000

Source data: dataset received on 4/11/2005, analysis data M1M21922

Table C5: Report on Test for Positive Dose-Tumor Trends in Female Mice

Organ Name	Tumor Name	CTR1	CTR2	LOW	MED	HIGH	P-Value (Exact Method)
ADRENAL CORTEX	B-SUBCAPSULAR CELL ADENOMA	0	2	2	1	2	0.2812
COLON	M-ADENOCARCINOMA	0	0	0	0	1	0.1946
DUODENUM	M-ADENOCARCINOMA	0	0	1	0	0	0.5738
ADRENAL MEDULLA	B-BENIGN PHEOCHROMOCYTOMA	0	0	1	1	0	0.5265
HARDERIAN GLAND	B-ADENOMA	2	1	2	3	3	0.2568
JEJUNUM	M-ADENOCARCINOMA	0	1	0	0	0	1.0000
KIDNEYS	M-TRANSITIONAL CELL CARCINOMA	0	0	1	0	0	0.5806
LIVER	B-HEPATOCELLULAR ADENOMA	1	1	0	0	0	1.0000
LIVER	B-HEMANGIOMA	0	1	0	1	0	0.6139
LIVER	M-HEPATOCELLULAR CARCINOMA	0	0	1	2	0	0.4492
LIVER	M-HEPATOBLASTOMA	0	0	0	0	1	0.2062
LUNG	B-BRONCHIOLO-ALVEOLAR ADENOMA	8	6	2	7	3	0.8087
LUNG	M-BRONCHIOLO-ALVEOLAR CARCINOM	3	5	8	3	5	0.5136
MAMMARY GLAND	M-ADENOCARCINOMA	2	2	0	1	1	0.7267
MAMMARY GLAND	M-ADENOACANTHOMA	0	1	0	0	0	1.0000
MULTISYSTEMIC	M-SCHWANNOMA	0	0	2	1	1	0.2552
MULTISYSTEMIC	M-MESOTHELIOMA	0	2	0	1	0	0.7950
MULTISYSTEMIC	M-HEMANGIOSARCOMA	5	3	2	1	1	0.9579
MULTISYSTEMIC	M-FIBROSARCOMA	1	0	0	1	1	0.2730
MULTISYSTEMIC	M-LEIOMYOSARCOMA	0	0	1	1	1	0.1847
MULTISYSTEMIC	M-HISTIOCYTIC SARCOMA	9	4	5	7	5	0.6193
MULTISYSTEMIC	M-LARGE GRANULAR LYMPHOCYTE LE	1	1	1	0	1	0.5749
MULTISYSTEMIC	M-LYMPHOMA	11	11	9	10	11	0.5421
MULTISYSTEMIC	M-OSTEOSARCOMA	1	0	0	1	0	0.6339
MULTISYSTEMIC	M-PLASMA CELL LYMPHOMA	0	1	0	0	0	1.0000
MULTISYSTEMIC	M-UNDIFFERENTIATED SARCOMA	0	0	1	0	0	0.5950
OVARIES	B-CYSTADENOMA	0	3	7	0	2	0.6773
OVARIES	B-LUTEOMA	2	1	0	2	1	0.5215
OVARIES	B-CYSTADENOMA, PAPILLARY, BILA	0	0	0	0	2	0.0449
OVARIES	M-CYSTADENOCARCINOMA	0	0	1	0	1	0.4427
OVARIES	B-BENIGN THECOMA	0	0	1	1	0	0.5364
OVARIES	B-GRANULOSA CELL TUMOR	1	0	0	0	0	1.0000
PANCREAS	B-ISLET CELL ADENOMA	0	0	1	0	0	0.5806
PARATHYROIDS	B-ADENOMA	0	0	0	1	0	0.7000
PITUITARY	B-ADENOMA	7	0	1	0	5	0.2688
PITUITARY	M-CARCINOMA	0	0	1	0	0	0.6357
SKIN	M-CARCINOMA, BASAL CELL	0	1	0	0	0	1.0000
SQUAMOUS STOMACH	B-SQUAMOUS CELL, PAPILLOMA	0	1	0	0	2	0.0978
STOMACH	B-ADENOMATOUS POLYP	1	0	0	0	0	1.0000

TONGUE	B-SQUAMOUS CELL PAPILLOMA	0	0	0	0	1	0.2023
UTERUS	M-ADENOCARCINOMA	0	1	2	2	0	0.6878
UTERUS	M-STROMAL CELL SARCOMA	4	3	2	3	2	0.7212
UTERUS	B-ENDOMETRIAL STROMAL POLYP	10	7	4	13	5	0.7220
UTERUS	B-LEIOMYOMA	1	0	1	2	1	0.3496
VAGINA	B-STROMAL POLYP	2	1	1	3	0	0.7188
BONE/JOINT	M-SQUAMOUS CELL CARCINOMA	1	0	0	0	0	1.0000

Source data: dataset received on 7/30/2006, analysis data M1F21992

Appears This Way
On Original

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

/s/

Moh-Jee Ng
1/10/2007 11:48:30 AM
BIOMETRICS

Karl Lin
1/10/2007 02:17:20 PM
BIOMETRICS
Concur with review



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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number: 21-992
Drug Name: DVS SR
Indication(s): Major Depression Disorder (MDD)
Applicant: Wyeth Pharmaceuticals
Date(s): December 22, 2005
Review Priority: Standard

Biometrics Division: Biometrics I (HFD-710)
Statistical Reviewer: Fanhui Kong
Concurring Reviewers: Peiling Yang, Kooros Mahjoob

Medical Division: Division of Psychiatry Products
Clinical Team: Robert Levin, Thomas Laughren
Project Manager: Renmeet, Gujral

Keywords: ANCOVA, MMRM, LOCF, family-wise error (FEW), Dunnett's Method, ETRANK, Major Depression Disorder.

Table of Contents

1. EXECUTIVE SUMMARY	3
1.1 CONCLUSIONS AND RECOMMENDATIONS	3
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES	3
1.3 STATISTICAL ISSUES AND FINDINGS.....	4
1.3.1 Studies 309, 317, 320.....	4
1.3.2 Studies 306, 308.....	4
2. INTRODUCTION	5
2.1 OVERVIEW	5
2.2 DATA SOURCES	6
3. STATISTICAL EVALUATION	6
3.1 EVALUATION OF EFFICACY	6
3.1.1 Studies 309, 317, 320.....	6
3.1.1.1 Dispositions.....	7
3.1.1.2 Demographic Characteristics.....	7
3.1.1.3 Patient Discontinuation	11
3.1.1.4 Statistical Issues and Results.....	11
3.1.2 Studies 306, 308.....	15
3.1.2.1 Dispositions.....	15
3.1.2.2 Demographic Characteristics	15
3.1.2.3 Patient Discontinuation	18
3.1.2.4 Statistical Issues and Results.....	18
3.2 EVALUATION OF SAFETY	21
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS.....	21
4.1 GENDER, RACE AND AGE	21
4.1.1 Studies 309, 317, 320.....	21
4.1.2 Studies 306, 308.....	22
4.2 OTHER SPECIAL/SUBGROUP POPULATIONS.....	23
5. SUMMARY AND CONCLUSIONS	24
5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE	24
5.1.1 Studies 309, 317, 320.....	24
5.1.2 Studies 306, 308.....	24
5.2 CONCLUSIONS AND RECOMMENDATIONS	25

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

In this submission, the sponsor conducted 7 short-term DVS SR studies to evaluate the efficacy of DVS SR in doses of 100, 200 or 400 mg/day in the treatment of Major Depression Disorder (MDD) in adult outpatients. Five studies (Studies 306, 308, 309, 317 and 320) were evaluated in this review. Two studies (Studies 306 and 308) provided evidence supporting the effectiveness of DVS SR in the treatment of adult outpatients with MDD with respect to the primary endpoint and other two studies (Studies 309 and 320) appeared to provide supportive information for such a claim. Study 317 was negative.

The primary efficacy analysis using LOCF in Studies 309, 317 and 320 did not support the claim of the effectiveness of DVS SR in the treatment of MDD. According to my evaluation, the non-significance might be resulted from unbalanced early dropouts due to adverse events. Removing these patients resulted in statistical significance using LOCF analyses in Studies 309 and 320. In addition, OC, MMRM and ETRANK analyses in these two studies also seemed to support the claim of the effectiveness of DVS SR in the treatment of MDD by reducing the primary outcome, namely HAM-D₁₇ total score. However, due to unbalanced early dropout caused by adverse events, the risk benefit ratio should be carefully considered before giving a definite conclusion. The claim on the key secondary endpoint was not supported by the data in these studies.

The analysis results in Studies 306 and 308 supported the efficacy claim of DVS SR in the treatment of MDD in adult patients. In Study 308, for the primary endpoint, both 200 mg and 400 mg dose groups of DVS SR met the significance criterion. The significance results were supported by nonparametric methods as well as other pre-specified analyses (OC, MMRM, and ETRANK). The key secondary endpoint in both dose groups also met the statistical significance requirement. In Study 306, for primary endpoint, DVS SR dose groups of 100 mg and 400 mg met the statistical significance criterion but DVS SR 200 mg dose group did not. For the key secondary endpoint, the results in the dose groups of 100 mg and 400 mg seemed supportive.

1.2 Brief Overview of Clinical Studies

Seven short-term DVS SR studies were submitted for the evaluation of the efficacy of DVS SR in doses of 100, 200 or 400 mg/day in the treatment of MDD in adult outpatients. There were 3 fixed-dose studies (Studies 223, 306, and 308) and 4 flexible-dose studies (Studies 304, 309, 317, and 320). All the studies were multicenter, randomized, double-blind, placebo-controlled, parallel-group studies in adults with MDD, with a double-blind treatment period of 8 weeks. The primary objective was to evaluate the efficacy and safety of DVS SR vs. placebo in adult outpatients with MDD. Studies 304 and 223 showed no effect of improving the primary endpoint so they are not reviewed here. The change from baseline in HAM-D₁₇ total score was the primary endpoint and CGI-I score was the key secondary endpoint.

After the screening period, subjects were treated during a double-blind period with doses of DVS SR ranging from 100 to 400 mg/day for 8 weeks. A total of 1211 subjects in the 7 studies were treated with DVS SR, 803 with placebo, and 244 with the reference treatment group venlafaxine ER (Ven ER). The ITT population was the primary efficacy analysis population. It included 1186 subjects in the DVS SR groups, 797 subjects in the placebo group, and 242 subjects in the Ven ER group. The majority of the

patients were white and the average age was 42 at baseline. The population consisted of 38% male and 62% female patients.

These studies were conducted between April 2004 and May 2005 in US, Europe, South Africa and Taiwan. Flexible-dose Studies 309 and 317 contained Ven ER as a reference in addition to placebo, while Study 320 only had placebo as the comparator. Fixed-dose Study 306 had four treatment arms: DVS SR 100 mg, DVS SR 200 mg, DVS SR 400 mg and placebo while fixed-dose Study 308 had only three treatment arms: DVS SR 200 mg, DVS SR 400 mg and placebo.

1.3 Statistical Issues and Findings

1.3.1 Studies 309, 317, 320

The primary efficacy analysis of LOCF was performed using ANCOVA on the change from baseline of the HAM-D₁₇ total score. The sequential testing procedure was used to control the overall Type I error rate in order to claim the efficacy on the key secondary endpoint. The primary analysis of LOCF in these studies did not support the claim of the effectiveness of DVS SR in the treatment of MDD. According to my evaluation, the non-significance might be resulted from unbalanced early dropouts due to adverse events. Removing these patients resulted in statistically significant results using LOCF analysis in Studies 309 and 320. In addition, OC, MMRM and ETRANK analyses also seemed to provide support for the efficacy of DVS SR in the reduction of the HAM-D₁₇ total score among MDD patients. However, since the unbalanced early dropouts were mainly caused by adverse events, the risk benefit ratio should be carefully considered before giving a definite conclusion. The claim on the key secondary endpoint was not supported by the data of these studies.

1.3.2 Studies 306, 308

The primary efficacy analyses of LOCF were conducted on the change from baseline of the HAM-D₁₇ total score using ANCOVA. Significance was adjusted by Dunnett's method. The key secondary efficacy variable was the CGI-I score. Closed testing procedures were used to compare the dose groups of DVS SR with placebo based on the primary efficacy variable in order to control the overall type I error rate.

The analysis results supported the efficacy claim of DVS SR on the primary endpoint in the treatment of MDD in both studies. In study 306, DVS SR dose groups of 100 mg and 400 mg met the statistical significance criterion but DVS SR 200 mg dose group did not. In study 308, both dose groups of DVS SR met the significance criterion. The significance results were supported by nonparametric methods as well as other pre-specified analyses of OC, MMRM and ETRANK.

For the secondary endpoint, statistical significance was achieved for both 200 mg and 400 mg dose groups in Study 308. However, in Study 306, the procedure that the sponsor specified and used may not protect the overall family wise error rate for testing the primary and secondary endpoints. Thus, strictly speaking, the statistical significance of the key secondary endpoint for 100 mg and 400 mg dose groups is difficult to conclude. However, the very small nominal p-values of the key secondary endpoint for these two doses seemed supportive.

2. INTRODUCTION

2.1 Overview

In this submission, 7 short-term DVS SR studies were included for the evaluation of the efficacy of DVS SR in doses of 100, 200 or 400 mg/day in the treatment of MDD in adult outpatients (Table 2.1). Three of the studies were fixed-dose studies (studies 223, 306, and 308) and 4 were flexible-dose studies (studies 304, 309, 317, and 320). All of these studies were multicenter, randomized, double-blind, placebo-controlled, parallel-group studies in adults with MDD, with a double-blind treatment period of 8 weeks.

Table 2.1: Short-Term Studies of DVS SR Efficacy in Patients with MDD

Study Number	CSR Number	DVS SR Dose (mg/day) ^a	Comparator(s)
Fixed Dose Studies			
0600D3-223-FR/PL/US/ZA	CSR-49148	200 or 400	Placebo
3151A1-306-US	CSR-57298	100, 200, or 400	Placebo
3151A1-308-EU/WW	CSR-57406	200 or 400	Placebo
Flexible Dose Studies			
3151A1-304-US	CSR-54022	100 to 200	Placebo
3151A1-309-EU	CSR-57536	200 to 400	Placebo, Ven ER ^b
3151A1-317-US	CSR-58757	200 to 400	Placebo, Ven ER ^c
3151A1-320-US	CSR-58759	200 to 400	Placebo

CSR=clinical study report; EU=Europe; FR=France; PL=Poland; US=United States; Ven ER=venlafaxine extended-release formulation; WW=worldwide; ZA=South Africa.

Note: This table provides the full study numbers, including the project prefixes and country abbreviations for the suffixes. Throughout this efficacy summary, the studies will be referred to without the prefixes and suffixes. The MDD studies were conducted in the United States, Europe, South Africa, and Taiwan.

a: The double-blind period was 8 weeks in all 7 studies, followed by a 1-week (studies 223 and 304) or 1- to 2-week taper period (studies 306, 308, 309, 317, and 320).

b: Venlafaxine ER dose 75 to 150 mg/day.

c: Venlafaxine ER dose 150 to 225 mg/day.

Source: Page 8 of sponsor's Summary of Clinical Efficacy

In all of these studies, after the screening period, subjects were treated during a double-blind period with doses of DVS SR ranging from 100 to 400 mg/day. A total of 1211 subjects in the 7 studies were treated with DVS SR, 803 with placebo, and 244 with Ven ER. The primary efficacy analysis population was the ITT population. The ITT population included 1186 subjects in the DVS SR groups, 797 subjects in the placebo group, and 242 subjects in the Ven ER group. The numbers of subjects in all studies are given in Table 2.2.

Appears This Way
On Original

Table 2.2: Sample Size in all the Studies of DVS SR Efficacy in Patients with MDD: ITT Population

Study Type	Placebo	DVS SR 100 mg/day	DVS SR 200 mg/day	DVS SR 400 mg/day	DVS SR ^a	Ven ER ^b	Total
Fixed-Dose Studies							
223	78	--	63	72	--	--	213
306	118	114	116	113	--	--	461
308	124	--	121	124	--	--	369
Flexible-Dose Studies							
304	114	--	--	--	120	-	234
309	120	--	--	--	116	127	363
317	125	--	--	--	110	115	350
320	118	--	--	--	117	-	235
Total (All Studies)	797	114	300	309	463	242	2225

MDD = major depressive disorder; Ven ER = venlafaxine extended-release formulation.

a: DVS SR doses in study 304 were 100 to 200 mg/day and in studies 309, 317, and 320 were 200 to 400 mg/day.

b: Venlafaxine ER doses in study 309 were 75 to 150 mg/day and in study 317 were 150 to 225 mg/day.

Source: Page 9 of sponsor's Summary of Clinical Efficacy

The majority of the patients were white and the average age was 42 at baseline. Patients were randomized into the corresponding treatment arms with the sample sizes as indicated in Table 2.1. The double blind treatment period was 8 weeks. The change from baseline in HAM-D₁₇ total score was the primary endpoint and CGI-I score was the key secondary endpoint. The tests were two sided and the overall significance level for each study was $\alpha=0.05$, and the stepwise testing procedure was used to protect the overall type 1 error rate.

Studies 304 and 223 showed no significant results of the treatment in the improvement of the primary endpoint so they are not reviewed here.

2.2 Data Sources

The electronic study report and electronic SAS transport data sets for the studies are provided in <\\Cdsub1\evsprod\n021992\0000\m5>.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Studies 309, 317, 320

Studies 309, 317 and 320 were the flexible dose studies in which subjects were randomly assigned to receive DVS SR 200 to 400 mg, placebo, or a comparative treatment Ven ER 75 to 150 mg (Study 320 did not have the ven ER arm). These studies were conducted between April 2004 and May 2005 (April 2004 to March 2005 for Study 309, March 2005 to May 2005 for Study 317 and July 2004 to May 2005 for Study 320) in US, Europe, South Africa and Taiwan. The primary objective was to compare the efficacy and safety of DVS SR with placebo in adult outpatients with MDD. Studies 309 and 317 contained Ven ER as a reference in addition to placebo, while Study 320 only had placebo as a comparator.

3.1.1.1 Dispositions

In Study 309, a total of 369 subjects were enrolled in the study, with 118, 128, and 123 subjects randomly assigned to receive DVS SR 200 to 400 mg, Ven ER 75 to 150 mg, and placebo, respectively. Of the subjects enrolled in the study, 364 subjects received at least 1 dose of test article and were included in the safety population, and 363 were included in the ITT population.

In Study 317, a total of 369 subjects were enrolled in the study, with 121, 121, and 127 subjects randomly assigned to receive DVS SR, Ven ER, and placebo, respectively. Thirteen (13) subjects had no data after baseline and were excluded from safety and efficacy analyses. The remaining 356 subjects were included in safety population. The ITT population had 350 subjects.

In Study 320, a total of 244 subjects were enrolled in the study: 123 were randomly assigned to receive DVS SR and 121 were randomly assigned to receive placebo. Nine (9) subjects had no data after baseline and were excluded from safety and efficacy population. Both safety population and ITT population had 235 subjects.

3.1.1.2 Demographic Characteristics

The patient baseline demographic characteristics appear in Tables 3.1 to 3.3 for these three studies. There appeared to be little difference between the means of these demographic variables of the treatment groups in the demographic characteristics. The Placebo group tended to have larger deviation in the duration of the current episode of depression in Studies 317 and 320. The baseline HAM-D₁₇ total scores, baseline CGI-Severity scores were compatible between the treatment groups in these studies. The majority of the patients were white and females. The average age was 45 in Study 309, 41 in Study 317 and 38 in Study 320.

*Appears This Way
On Original*

Table 3.1 Baseline Demographic Characteristics for Study 309--ITT Population

Characteristic	Placebo (n=120)	DVS SR 200 to 400 mg (n=117)	Venlafaxine ER 75 to 150 mg (n=127)
Age, years			
Mean \pm SD	45.4 \pm 12.0	43.9 \pm 12.3	45.9 \pm 12.0
Range	21.0 to 72.0	18.0 to 73.0	23.0 to 70.0
Median	47.0	45.0	47.0
Sex, n (%)			
Female	80 (67)	84 (72)	92 (72)
Male	40 (33)	33 (28)	35 (28)
Ethnic origin, n (%)			
Black	0	1 (<1)	0
White	120 (100)	115 (98)	125 (98)
Asian	0	0	1 (<1)
Other	0	1 (<1)	1 (<1)
Height, cm			
Mean \pm SD	168.5 \pm 8.4	168.3 \pm 9.4	167.5 \pm 7.0
Range	149.0 to 187.0	150.0 to 198.0	154.0 to 186.0
Median	167.0	166.5	167.0
Weight, kg			
Mean \pm SD	73.2 \pm 16.1	72.1 \pm 16.9	72.0 \pm 14.9
Range	47.6 to 132.5	42.1 to 112.5	42.0 to 130.6
Median	69.6	70.0	70.3
Duration of current episode, months			
Mean \pm SD	6.5 \pm 11.8	6.9 \pm 12.7	7.6 \pm 16.4
Range	1.1 to 108.7	1.0 to 93.7	0.8 to 140.1
Median	2.8	2.9	3.0

Source: Page 52 of sponsor's Study Report of CSR-57536

**Appears This Way
On Original**

Table 3.2 Baseline Demographic Characteristics for Study 317--ITT Population

Characteristic	Placebo (n=125)	DVS SR 200-400 mg (n=110)	ven ER 150-225mg (n=115)
Age (year)			
Mean	39.5	41.3	41.5
SD	11.4	11.9	11.6
Min-max	18.0-64.0	18.0-68.0	21.0-73.0
Median	40.0	43.5	41.0
Sex, n (%)			
Female	80 (64)	70 (64)	80 (70)
Male	45 (36)	40 (36)	35 (30)
Ethnic origin, n (%)			
Black	27 (22)	12 (11)	19 (17)
Hispanic	8 (6)	4 (4)	3 (3)
Native American	1 (<1)	0	0
Oriental (Asian)	3 (2)	2 (2)	2 (2)
Other: Brazil and Portugal		1 (<1)	0
Other: Alaskan native	1 (<1)	0	0
Other: Indian	1 (<1)	1 (<1)	0
Other: Middle Eastern	1 (<1)	0	0
Other: Russian	0	0	1 (<1)
Other: Slavic	1 (<1)	0	0
Other: Spanish	0	0	1 (<1)
Other: biracial	0	1 (<1)	0
Other: Persian	0	1 (<1)	0
White	82 (66)	88 (80)	89 (77)
Height (cm)			
Mean	168.2	168.2	167.5
Standard deviation	9.9	9.9	9.6
Min-max	144.8-193.0	144.8-198.1	144.8-195.6
Median	167.6	167.6	167.6
Weight (kg)			
Mean	85.4	82.7	82.8
Standard deviation	21.3	22.9	21.8
Min-max	45.9-158.4	44.9-161.2	48.1-177.7
Median	83.1	79.0	78.5
Duration of current episode (months)			
Mean	19.8	12.7	16.4
Standard deviation	48.3	18.8	39.9
Min-max	1.0-475.1	1.0-166.6	1.0-324.7
Median	8.6	7.4	7.1

Source: Page 54-55 of sponsor's Study Report of CSR-58757

Appears This Way
On Original

Table 3.3 Baseline Demographic Characteristics for Study 320--ITT Population

Characteristic	DVS SR		
	Placebo (n = 118)	200-400 mg (n = 117)	Total (n = 235)
Age, year			
Mean	38.7	37.5	38.1
Standard Deviation	12.2	11.5	11.9
Min, Max	18.0, 74.0	19.0, 70.0	18.0, 74.0
Median	36.0	38.0	37.0
Sex, n (%)			
Female	81 (69)	72 (62)	153 (65)
Male	37 (31)	45 (38)	82 (35)
Ethnic Origin, n (%)			
Arabic	1 (<1)		1 (<1)
Asian	3 (3)	2 (2)	5 (2)
Black	16 (14)	22 (19)	38 (16)
Hispanic	4 (3)	11 (9)	15 (6)
Native American	2 (2)	1 (<1)	3 (1)
Other	1 (<1)		1 (<1)
Other: Eastern Europe.	1 (<1)		1 (<1)
Other: African American, American Indian, Asian/Pacific Islander, White		1 (<1)	1 (<1)
Other: Brazilian.	1 (<1)		1 (<1)
Other: White/Asian Pacific-Islander	1 (<1)		1 (<1)
Other: Malaysian.		1 (<1)	1 (<1)
Other: Pacific Islander.	1 (<1)		1 (<1)
White	87 (74)	79 (68)	166 (71)
Height, cm			
Mean	168.6	169.0	168.8
Standard Deviation	9.1	9.3	9.2
Min, Max	148.6, 208.3	149.4, 191.8	148.6, 208.3
Median	167.6	167.6	167.6
Weight, kg			
Mean	81.7	84.0	82.8
Standard Deviation	21.1	20.7	20.9
Min, Max	48.1, 145.3	47.7, 145.7	47.7, 145.7
Median	76.9	81.9	79.9
Duration of current episode, months			
Mean	27.2	19.3	23.3
Standard Deviation	62.4	24.7	47.6
Min, Max	1.0, 588.4	1.3, 157.3	1.0, 588.4
Median	10.1	9.2	9.3

Source: Page 48-49 of sponsor's Study Report of CSR-58759

3.1.1.3 Patient Discontinuation

In Study 309, 54 (15%) subjects discontinued treatment during the double-blind period: 13 (11%), 25 (21%), and 16 (13%) in the placebo, DVS SR, and Ven ER groups, respectively. In Study 317, 77 (22%) subjects discontinued treatment during the double-blind period: 17 (14%) in placebo, 35 (31%) in DVS SR, and 25 (21%) in Ven ER. In Study 320, 44(19%) subjects discontinued treatment during the double-blind period: 15 (13%) in placebo and 29 (25%) in DVS SR. Table 3.4 summarizes the number of subjects who discontinued treatment by the primary reasons for withdrawal in each treatment group.

Table 3.4 Number (%) of Subjects Who Discontinued Treatment During the Double-Blind Period by Primary Reason for Withdrawal

Study Reason of withdrawal	Placebo n (%)	DVS SR 200 to 400 mg n (%)	Venlafaxine ER n (%)
Study 309	(n=120)	(n=117)	(n=127)
Total withdrawal	13 (11)	25 (21)	16 (13)
Adverse event	1 (<1)	19 (16)	7 (6)
Failed to return	1 (<1)	1 (<1)	1 (<1)
Other event	1 (<1)	1 (<1)	3 (2)
Subject request unrelated to study	1 (<1)	1 (<1)	1 (<1)
Protocol violation	3 (3)	1 (<1)	1 (<1)
Unsatisfactory response -- efficacy	6 (5)	2 (2)	3 (2)
Study 317	(n=125)	(n=114)	(n=117)
Total withdrawal	17 (14)	35 (31)	25 (21)
Adverse event	6 (5)	19 (17)	10 (9)
Failed to return	5 (4)	9 (8)	6 (5)
Other event	3 (2)	4 (4)	2 (2)
Subject request unrelated to study	0	1 (<1)	2 (2)
Protocol violation	0	2 (2)	3 (3)
Unsatisfactory response -- efficacy	3 (2)	0	2 (2)
Study 320	(n=118)	(n=117)	
Total withdrawal	15 (13)	29 (25)	
Adverse event	4 (3)	15 (13)	
Failed to return	6 (5)	8 (7)	
Other event	0	1 (<1)	
Subject request unrelated to study	2 (2)	1 (<1)	
Protocol violation	2 (2)	3 (3)	
Unsatisfactory response -- efficacy	1 (<1)	1 (<1)	

Source: Table 8.1.1-1 of sponsor's Study Report of Study CSR-57536, Table 8.1.1-1 of Study Report of Study CSR-58757 and Table 8.1.1-1 of Study Report of Study CSR-58759.

3.1.1.4 Statistical Issues and Results

Missing items for observations on the scales were handled as follows: if more than 50% of the items were missing from one of these scales, then the total score for this scale was not used in the analysis. If 50% or fewer items were missing from one of these scales, then the average of the available items was multiplied by the total number of items to obtain an inferred score.

The primary efficacy analysis was performed on the change from baseline of the HAM-D₁₇ total score at the final on-therapy evaluation in the ITT population and the treatment effect was tested using the analysis of covariance (ANCOVA) with treatment and site as factors and the baseline HAM-D₁₇ total score as the covariate. The key secondary endpoint, the CGI-I score, was analyzed using the analysis of variance (ANOVA) with treatment and site as factors. The sequential testing procedure was planned in SAP to control the overall family wise Type I error rate. This procedure requires the primary endpoint to be significant before the key secondary endpoint is tested. Statistical significance was tested at a nominal significance level of 0.05 (two-sided). LOCF was used as the primary analysis for the missing observations of the dropout patients. No interim analyses were performed.

According to SAP, a treatment-by-site interaction term was to be tested in the primary and key secondary efficacy analysis models to explore the possibility of qualitative or quantitative treatment-by-site interaction. If the interaction was significant ($p \leq 0.10$), an assessment of the magnitude and direction of the interaction term was to be made. At the same time, the test of homoscedasticity was to be performed with the Levene test. A covariate and treatment interaction term was added to the primary model to test parallelism. Normality was to be tested using the Shapiro-Wilk's test. If the normality assumptions of ANCOVA were not met, an ANOVA or a nonparametric ANCOVA based on ranks was to be performed.

Table 3.5: Statistical Comparisons (p-Values) Between Treatments and Placebo for the Primary and Key Secondary Efficacy Variables in Flexible-Dose Studies 309, 317, and 320 at Week 8

Efficacy Variable	Study 309 ^a		Study 317 ^a		Study 320 ^a
	DVS SR ^b	Ven ER ^b	DVS SR ^b	Ven ER ^b	DVS SR ^b
Primary Variable: HAM-D₁₇ total score					
Final on-therapy evaluation (LOCF analysis)	0.381	0.171	0.488	0.005	0.078
Observed-case analysis	<0.001	0.027	0.173	0.008	0.008
Mixed-effect model ^c	0.011	0.153	0.074	<0.001	0.012
ETRANK ^d	0.028	0.253	0.353	<0.001	<0.001
Key Secondary Variable: CGI-I score					
Final on-therapy evaluation (LOCF analysis)	0.404	0.107	0.604	0.011	0.037
Observed-case analysis	<0.001	0.014	0.141	0.010	0.006

a: Significant differences between active treatment and placebo are shown in bold.

b: DVS SR doses in Studies 309, 317, and 320 were 200 to 400 mg/day. Ven ER doses in Study 309 were 75 to 150 mg/day and in Study 317 were 150 to 225 mg/day.

c: Mixed-effect model with AR(1) covariance structure.

d: ETRANK analysis: full data analysis with Entsuah score empirical significance levels (p-values).

Source: Tables 1.4.1.2-1, 1.4.1.2-2, 1.4.2.2-1, 1.4.2.2-2 in sponsor's Summary of Clinical Efficacy.

In longitudinal studies, patient dropout raises concerns on the interpretation of efficacy results. In this submission, LOCF was pre-specified as the procedure for primary analysis. Yet it is difficult to evaluate the impact of this procedure on efficacy results. If the dropout rate is low, say below 5%, such effect may

be negligible. For the dropout rates in these studies (16% for Study 309, 22% for Study 317 and 18% for Study 320), the impact of LOCF procedure on efficacy result is not clear. In general, if the mean outcome measure is stable over the whole study period, the impact of LOCF procedure on the efficacy results may be limited. Otherwise, it could dramatically affect the significance results.

The results were given by the sponsor and they are depicted in Table 3.5. No statistical significance was observed in the primary analysis of LOCF on the primary endpoint between DVS SR and placebo at final on-therapy evaluation in Studies 309, 317 and 320.

Table 3.6: Treatment Effects in Treatment Groups for the Primary Efficacy Variable in Flexible-Dose Studies 309, 317, and 320 at Week 8

Mean Change in HAM-D ₁₇ total score	Study 309			Study 317			Study 320	
	Placebo	DVS SR	Ven ER	Placebo	DVS SR	Ven ER	Placebo	DVS SR
Final on-therapy evaluation, LOCF (SE)	-12.5 (0.71)	-13.4 (0.72)	-13.8 (0.69)	-9.78 (0.73)	-10.5 (0.78)	-12.6 (0.76)	-7.48 (0.66)	-9.08 (0.67)
Observed-case analysis (SE)	-12.8 (0.69)	-16.4 (0.70)	-14.9 (0.65)	-10.7 (0.85)	-12.4 (0.98)	-13.8 (0.89)	-7.87 (0.74)	-10.7 (0.81)

Source: Tables 1.4.1.2-1 in sponsor’s Summary of Clinical Efficacy.

To explain non-significance of the primary analysis results in Study 309, the sponsor pointed out that: “the lack of significant difference in efficacy between active treatments and placebo appears to result from a clinically important placebo effect.” They further pointed out on Page 79 of the Final Report of Study 309:

The absence of statistically significant differences from baseline in the pairwise comparisons for DVS SR and venlafaxine ER against placebo may be explained by a high response rate in the placebo group. Additionally, the separation of the active treatment from placebo in the observed cases analysis suggests that the early study withdrawals had a large effect on the LOCF analysis. The important placebo effect was investigated: 3 sites mainly contributed to this high placebo response. There was a significant treatment by site interaction in this study, with sites in South Africa and the Baltic nations, having a considerable drop in the HAM-D₁₇ score from baseline for placebo treated subjects, while centers in France and Belgium had a greater drop for DVS SR treated subjects than in other sites.

The sponsor did not give specific figures of how these centers affected the significance results. While conducting detailed investigation of the impact of these dropouts on efficacy results, the reviewer found that there were large differences in early dropout among different treatment groups, especially the patient who had only one post-randomization observation. These patients dropped out before having much improvement on efficacy measure. Such early dropout in treatment groups was quite unbalanced. In Study 309, there were 21 such dropouts in total, 14 were in DVS SR group, 7 were in the Ven ER group and 0 was in the placebo group. After these patients were removed, the LOCF analysis gave a p-value of 0.0065 for DVS SR and 0.015 for Ven ER when compared to placebo. In Study 320, there were a total of 16 such dropout patients, 12 were in the DVS SR group and 4 were in the placebo group. After these patients were removed, the LOCF analysis gave a p-value of 0.04. These results suggested that in these studies, the non-significance of the LOCF analysis was caused by the unbalance of the very early dropout between the treatment and placebo groups. However, further investigation indicated that these early dropout were

mainly caused by “Adverse Event”. In Study 309, 9 out of 14 early dropouts in the DVS SR group were caused by “Adverse Event”, 1 was by “Protocol Violation” and 4 were by unknown reason of dropout. In Ven ER group, 4 out of the 7 dropouts were caused by “Adverse Event”, 1 was by “Protocol Violation”, 1 was by “Other Reason” and 1 was unknown. In Study 320, 10 out of the 12 early dropouts in DVS SR group were caused by “Adverse Event”, 2 were unknown. For the 4 early dropouts in the placebo group of Study 320, 2 failed to return and 2 were unknown. Therefore, in addition to considering the impact of these early dropouts on the efficacy evaluation, one should also evaluate their impact on drug safety. This will be referred to medical officers.

The normality assumption for the primary endpoint the HAM-D₁₇ total score was rejected in these studies so such an assumption was not supported by data. The homoscedasticity was not tested. The sponsor believed that the sample sizes in these studies were sufficiently large for asymptotic normality theory to be valid. As sensitivity studies, nonparametric ANCOVA was performed, using the rank of the change from baseline of the HAM-D₁₇ total score as the response, with treatment and site as factors and ranked baseline score as the covariate. Nonparametric ANCOVA analysis for LOCF data gave *p*-values of 0.31, 0.43 and 0.055 for DVS SR in Studies 309, 317 and 320, respectively. The corresponding Wilcoxon rank sum test gave *p*-values of 0.39, 0.43 and 0.041. Such analyses do not depend on model assumptions. Therefore, the normality assumption in the ANCOVA seemed to have very limited effect on the significance results.

To further see how the patient dropout affected the significance results, three additional analyses were performed on the HAM-D₁₇ total score as specified in the SAP: the ANCOVA on observed-cases analysis (OC), mixed-effect model with repeated measurement analysis (MMRM) and ETRANK analysis at Week 8. In general, the OC analysis does not use the whole ITT population when there are patient dropouts. It eliminates the patients who dropped out before the end of the study. Among the dropouts, those due to “Adverse Event”, which happens more likely in the active treatment groups, and those due to the lack of treatment effect, which happens more likely in the placebo group make it hard to interpret the OC analysis results. One way to make up the lost information is to see how the outcome change for those who finished the study yet suffered the similar kind of adverse events. Nonetheless, positive result in OC analysis provides some indication of the effectiveness of the treatment.

The mixed-effect model with repeated measurement procedure, namely MMRM, uses ITT population and includes the information of all the observed outcome measurement in order to evaluate the treatment efficacy at the end of study. This procedure could give more reliable efficacy results if the patient dropout is non-informative, with dropout only depending on the observed outcome values, not on unobserved values. However, this assumption cannot be directly verified. Although only specified as the secondary analysis, positive result in MMRM provided supportive information to the effectiveness of the treatment.

ETRANK is a nonparametric procedure analyzing repeated measurements data with patient dropout. It provides tests of the treatment effect at the end of study by incorporating all available data through adjusting for withdrawal patterns, the proportion of withdrawal and the level of response prior to withdrawal using data dependent scoring schemes. Positive result in ETRANK provided supportive information to the effectiveness of the treatment.

In study 309, significant efficacy results were observed between DVS SR and placebo in the OC, MMRM, and ETRANK analyses at Week 8 ($p < 0.001$, $p = 0.011$, $p = 0.028$; respectively). As a reference, significant efficacy results were also observed between Ven ER and placebo in the OC analysis ($p = 0.027$), but not in the mixed effect model and ETRANK analyses. These results seemed to provide supportive information on the efficacy of DVS SR in the treatment of MDD on the primary efficacy

measure of the HAM-D₁₇ total score. However, because of the unbalanced early dropout caused by adverse events, the risk benefit ratio should be carefully evaluated.

None of these analyses gave positive efficacy results for DVS SR vs. placebo in Study 317. So this study didn't provide enough evidence to support to the claim of the effectiveness of DVS SR in the treatment of MDD.

In Study 320, significant efficacy results were observed for the DVS SR group versus placebo in OC, MMRM, and ETRANK analyses (p=0.008, 0.012, and <0.001; respectively). Again, these results seemed to provide supportive evidence for DVS SR to improve the HAM-D₁₇ total score in the treatment of MDD. However, because of the unbalanced early dropout caused by adverse events in treatment groups, the risk benefit ratio should be carefully evaluated.

The key secondary endpoint was not significant in these studies.

In conclusion, the protocol specified primary analysis using LOCF data in flexible dose Studies 309, 317 and 320 did not support the claim of the effectiveness of DVS SR in the treatment of MDD. The reviewer's investigation suggested that such non-significance could be the result of unbalanced early dropout caused by adverse events. Removing these patients changed the non-significant results in Studies 309 and 320 to be significant. In addition, OC, MMRM and ETRANK analyses in these two studies also provided supportive evidence for the effectiveness of DVS SR in the treatment of MDD in reducing the HAM-D₁₇ total score. However, due to the unbalanced early dropout caused by adverse events, the risk benefit ratio should be carefully evaluated before giving a definite conclusion.

3.1.2 Studies 306, 308

These were phase 3, multicenter, randomized, double blind, placebo-controlled, fixed-dose studies in which eligible patients were randomly assigned to either a treatment group or placebo and followed for approximately 8 weeks. These treatment arms were: DVS SR 100 mg, 200 mg, 400 mg and placebo in Study 306, DVS SR 200 mg, 400 mg and placebo in Study 308.

3.1.2.1 Dispositions

In Study 306, a total of 480 subjects were randomly assigned to four groups: 121 to placebo; 120 to DVS SR 100 mg; 120 to DVS SR 200 mg; and 119 to DVS SR 400 mg. Ten (10) subjects had no data after baseline. The remaining 470 subjects were included in all safety analyses. The ITT population included 461 subjects. This study was conducted in the United States from November 2003 to November 2004.

In Study 308, 375 subjects were randomly assigned to 3 treatment groups: 126 to placebo; 124 to DVS SR 200 mg; and 125 to DVS SR 400 mg. Two (2) subjects had no data after baseline. The remaining 373 subjects were included in the safety analyses. The ITT population included 369 subjects. This study was conducted in Europe and South Africa from November 2003 to September 2004.

3.1.2.2 Demographic Characteristics

The patient baseline demographic characteristics appear in Tables 3.6 to 3.7 for the two studies. There were no significant differences between treatment groups in the demographic characteristics or baseline

characteristics. The majority of the patients were white and females. The average age was 40 in Study 306, and 45 in Study 320.

Table 3.6 Demographic and Baseline Characteristics for Study 306--ITT Population

Characteristic	Placebo (n = 118)	DVS SR 100 mg (n = 114)	DVS SR 200 mg (n = 116)	DVS SR 400 mg (n = 113)
Age (year), n	118	114	116	113
Mean	40.0	40.4	40.7	39.0
SD	12.8	12.1	12.8	12.6
Min, max	18.0, 64.0	18.0, 72.0	18.0, 75.0	18.0, 72.0
Sex, n (%)				
Female	80 (68)	74 (65)	71 (61)	61 (54)
Male	38 (32)	40 (35)	45 (39)	52 (46)
Ethnic origin, n (%)				
Arabic	1 (<1)	1 (<1)	0 (0)	0 (0)
Asian	2 (2)	3 (3)	2 (2)	3 (3)
Black	11 (9)	11 (10)	16 (14)	10 (9)
Hispanic	18 (15)	10 (9)	19 (16)	14 (12)
Native American	0 (0)	0 (0)	2 (2)	0 (0)
Other	0 (0)	1 (<1)	2 (2)	2 (2)
White	86 (73)	88 (77)	75 (65)	84 (74)
Height (cm), n	116	114	115	113
Mean	168.3	168.1	168.2	170.4
SD	9.3	10.3	10.7	9.0
Min, max	149.9, 195.6	149.9, 194.3	149.9, 200.7	152.4, 194.8
Weight (kg), n	118	114	116	113
Mean	81.7	85.1	83.0	83.7
SD	20.3	20.2	22.7	21.1
Min, max	51.1, 149.4	46.3, 138.5	45.9, 213.4	51.3, 168.9
Duration of current episode (months), n	118	114	116	113
Mean	24.0	27.8	26.1	20.8
SD	51.2	58.3	55.1	29.3
Min, max	1.1, 466.1	1.1, 411.0	1.1, 480.0	1.0, 209.5

Source: Page 52 of sponsor's Study Report of CSR-57298

**Appears This Way
On Original**

Table 3.7 Demographic and Baseline Characteristics for Study 308--ITT Population

Characteristic	Placebo (n = 124)	DVS SR 200 mg (n = 121)	DVS SR 400 mg (n = 124)
Age, years			
Mean	46.7	45.4	43.4
SD	12.1	11.9	10.9
Min-max	19.0 - 73.0	19.0 - 69.0	18.0 - 69.0
Sex, n (%)			
Female	84 (68)	83 (69)	77 (62)
Male	40 (32)	38 (31)	47 (38)
Ethnic origin, n (%)			
Black	1 (<1)		
Other	1 (<1)	1 (<1)	2 (2)
White	122 (98)	120 (99)	122 (98)
Height, cm			
Mean	167.1	167.6	168.1
SD	8.0	8.5	8.7
Min-max	150.0 - 187.0	147.0 - 190.0	150.0 - 190.0
Weight, kg			
Mean	71.7	75.0	72.7
SD	15.8	17.0	16.3
Min-max	45.0 - 112.0	47.0 - 122.0	44.0 - 130.0
Duration of current episode, months			
Mean	7.1	8.5	7.1
SD	8.1	16.9	13.3
Min-max	1.0 - 51.3	1.1 - 152.0	1.2 - 120.7

Source: Page 52 of sponsor's Study Report of CSR-57406

Appears This Way
On Original

3.1.2.3 Patient Discontinuation

In Study 306, 110 (23%) subjects discontinued treatment during the double-blind period: 22 (18%) in placebo group, 27 (23%) in DVS SR 100 mg group, 26 (22%) in DVS SR 200 mg group, and 35 (30%) in DVS SR 400 mg group. In Study 308, 93 (25%) subjects discontinued treatment during the double-blind period: 27 (22%) in placebo group, 33 (27%) in DVS SR 200 mg group, and 33 (26%) in DVS SR 400 mg group. Table 3.8 summarizes the number of subjects who discontinued treatment by the primary reasons for withdrawal in each treatment group. In Study 306, treatment groups had a higher percentage of dropout due to adverse events. In Study 308, treatment groups had a high percentage of dropout due to adverse events and the placebo group had a high percentage of dropout due to unsatisfactory response.

Table 3.8 Number (%) of Subjects Who Discontinued Treatment During the Double-Blind Period by Primary Reason for Withdrawal

Study Reason	Placebo	DVS SR 100 mg	DVS SR 200 mg	DVS SR 400 mg
Study 306	(n=120)	(n=118)	(n=116)	(n=116)
Total	22 (18)	27 (23)	26 (22)	35 (30)
Adverse event	4 (3)	15 (13)	11 (9)	19 (16)
Failed to return	8 (7)	8 (7)	9 (8)	10 (9)
Other event	2 (2)	2 (2)	3 (3)	0
Subject request unrelated to study	2 (2)	1 (<1)	1 (<1)	2 (2)
Protocol violation	1 (<1)	1 (<1)	1 (<1)	2 (2)
Unsatisfactory response -- efficacy	5 (4)	0	1 (<1)	2 (2)
Study 308	(n=125)		(n=123)	(n=125)
Total	27 (24)		33 (27)	33 (26)
Adverse event	7 (6)		25 (20)	26 (21)
Failed to return	0		0	2 (2)
Other event	2 (2)		0	0
Subject request unrelated to study	2 (2)		1 (<1)	1 (<1)
Protocol violation	1 (<1)		3 (2)	2 (2)
Unsatisfactory response -- efficacy	15 (12)		4 (3)	2 (2)

Source: Table 8.1.1-1 of sponsor's Study Report of Study CSR-57298 and Table 8.1.1-1 of sponsor's Study Report of Study CSR-57406

3.1.2.4 Statistical Issues and Results

Just as in the flexible dose studies, the primary efficacy analyses in Studies 306 and 308 were also performed on the change from baseline in the HAM-D₁₇ total score at Week 8 in the ITT population with treatment effect tested using ANCOVA. The key secondary efficacy variable, the CGI-I score, was analyzed using ANOVA with treatment and site as factors. Statistical significance was tested at a nominal significance level of 0.05 (2-sided). LOCF was used as the primary analysis for the missing observations caused by patient dropout.

Study 306 had four treatment arms: DVS SR 100 mg, 200 mg, 400 mg and placebo. Study 308 had only three treatment arms: DVS SR 200 mg, 400 mg and placebo. They all had the primary endpoint of the HAM-D₁₇ total score and the key secondary endpoint the CGI-I score. These studies used different procedures to control the overall Type 1 error rate.

In Study 306, it was described in the SAP and Final Report (6.7.1.1 and 6.7.1.1.1):

A multiple-comparisons procedure, the Dunnett's procedure, was used to address the multiplicity issue associated with testing the 3 DVS SR treatment groups against placebo for the primary and the key secondary efficacy variables.

"A sequential testing strategy with a prespecified order of testing was used to control for multiplicity in the primary and key secondary efficacy variables. This was applied separately for each dose of DVS SR compared with placebo. If the comparison of DVS SR and placebo was significant at the ≤ 0.05 level for the primary efficacy variable (HAM-D₁₇), then the subsequent pairwise comparison of the key secondary efficacy variable (CGI-I) was made and was declared significant if the p-value was ≤ 0.05 . The key secondary efficacy variable, the CGI-I score, was considered significant only if the previous comparison (for the primary efficacy variable) was significant and the p-value for the key secondary efficacy variable was ≤ 0.05 . All other pairwise comparisons were considered exploratory."

This procedure may not control the family-wise error (FEW) rate. Here the sponsor proposed a hierarchical procedure in which both the primary and key secondary endpoints (each having multiple dose levels) could lead to regulatory claims. However, the procedure that the sponsor specified is not a "serial gatekeeping procedure" in straight sense, nor is it a "parallel gatekeeping procedure" such as those proposed by Dmitrienko *et al.* (2003) or Chen *et al.* (2005). It may not control the overall FEW type I error rate. In fact, if the null hypotheses regarding DVS SR 400 mg in HAM-D₁₇, namely H₁₃, DVS SR 100 mg and DVS SR 200 mg in CGI-I, namely H₂₁ and H₂₂, were true, while all the other null hypotheses were not true, with the treatment much better than the placebo, then the FWE will approximately be $\text{prob}(H_{13} \text{ or } H_{21} \text{ or } H_{22} \text{ is false})$, which could be larger than 0.05. A possible procedure that controls the overall family-wise type I error is to test all the null hypotheses with primary endpoint HAM-D₁₇ using the Dunnett procedure first. If all are significant, then test the null hypotheses with all the doses of secondary endpoint CGI-I and declare the dose to be significant if the corresponding test is significant.

In Study 308, it was described in the SAP and the Final Report (6.7.1.1):

The stepwise procedure was used to control the overall Type I error rate. Closed testing procedures were performed to compare the 2 doses (200 and 400 mg/day) of DVS SR with placebo based on the primary efficacy variable, the HAM-D₁₇ change from baseline. A general linear model with multiple contrast statements was used to calculate F-statistics for the global null hypotheses and all intersection hypotheses. The closure principle was used to determine which hypothesis should have been retained or rejected at $\alpha=0.05$. If a significant difference was detected for one or both doses of DVS SR, then a sequential testing method was applied to that dose(s) as follows. For 1 or both DVS SR dose group(s), if a significant difference to placebo on the primary efficacy variable was noted based on the closed testing procedure, the key secondary efficacy variables were tested at the 0.05 level to compare the DVS dose(s) to placebo.

For the same reason as given above, this procedure may not control the family wise error rate. If the sequential or hierarchical testing procedure is to be used, the key secondary endpoint should be tested only when **all** the primary null hypotheses are rejected.

The p-values adjusted by the Dunnett's method in primary analyses using LOCF data in 2 pivotal studies, 306 and 308, are depicted in Table 3.9. These results supported the efficacy claim of DVS SR in the

treatment of MDD. In Study 306, DVS SR dose groups of 100 mg and 400 mg was statistically significantly better than placebo in the reduction of the HAM-D₁₇ total score at the final on-therapy evaluation (p=0.004 and p=0.002, respectively). The DVS SR 200 mg dose group did not meet the significance requirement (p=0.076). The reviewer conducted the analyses and derived the same adjusted p-values. In Study 308, both dose groups of DVS SR (200 mg and 400 mg) were statistically significantly better than placebo in the reduction of the HAM-D₁₇ total score at the final on-therapy evaluation (p=0.002 and p=0.008, respectively). The reviewer did the same analysis and derived the corresponding adjusted p-values for the primary endpoint to be 0.004 and 0.015. The corresponding adjusted p-values for the key secondary endpoint of the CGI-I score became 0.008 and 0.052.

Table 3.9: Statistical Comparisons (p-values) between Treatment and Placebo for the Primary and Key Secondary Efficacy Variables in Fixed-Dose Studies 306 and 308^a at Week 8

Efficacy Variable	Study 306			Study 308	
	DVS SR 100 mg	DVS SR 200 mg	DVS SR 400 mg	DVS SR 200 mg	DVS SR 400 mg
Primary Variable: HAM-D₁₇ total score					
Final on-therapy evaluation (LOCF analysis)	0.004	0.076	0.002	0.002	0.008
Observed-case analysis	0.001	0.002	<0.001	<0.001	<0.001
Mixed-effect model	<0.001	0.004	<0.001	<0.001	<0.001
ETRANK	<0.001	<0.001	<0.001	<0.001	<0.001
Key Secondary Variable: CGI-I score					
Final on-therapy evaluation (LOCF analysis)	0.001	0.046	0.013	0.004	0.028
Observed-case analysis	<0.001	<0.001	<0.001	<0.001	<0.001

a: these p-values are the adjusted p-values using Dunnett's procedure. CGI-I=Clinical Global Impressions Scale-Improvement; HAM-D₁₇=Hamilton Rating Scale for Depression, 17-item; Ven ER= venlafaxine extended-release formulation.

Source: Table 4.2.1-1 of sponsor's Clinical Overview.

Table 3.10: Treatment Effect in Treatment Groups for the Primary Efficacy Variable in Fixed-Dose Studies 306, 308 at Week 8

Efficacy Variable	Study 306				Study 308		
	Placebo	DVS SR 100 mg	DVS SR 200 mg	DVS SR 400 mg	Placebo	DVS SR 200 mg	DVS SR 400 mg
Mean Change in HAM-D₁₇ total score							
Final on-therapy evaluation, LOCF (SE)	-7.7 (0.63)	-10.5 (0.66)	-9.6 (0.66)	-10.5 (0.67)	-9.3 (0.74)	-12.6 (0.75)	-12.1 (0.74)
Observed-case analysis (SE)	-8.0 (0.72)	-11.7 (0.73)	-11.6 (0.73)	-13.4 (0.76)	-11.0 (0.77)	-16.4 (0.81)	-15.2 (0.81)

Source: Tables 9.1-1 in sponsor's Study Report of csr-57298 and Tables 9.1-1 in sponsor's Study Report of csr-57406.

According to SAP, a treatment-by-site interaction term was to be tested in the primary and key secondary efficacy analysis models to explore the possibility of qualitative or quantitative treatment-by-site interaction. If the interaction was significant (p≤0.10), an assessment of the magnitude and direction of

the interaction term was to be made. At the same time, the test of homoscedasticity was to be performed with the Levene test. An interaction term using the covariate and treatment was added to the primary model to test parallelism. Normality was to be tested using the Shapiro-Wilk's test. If assumptions of ANCOVA were not met, an ANOVA or a nonparametric ANCOVA based on ranks was to be performed.

The normality assumption for the primary endpoint of the HAM-D₁₇ total score was rejected in these studies. The homoscedasticity was not tested. The sponsor believed that the sample sizes in these studies were sufficiently large for asymptotic normality theory to be valid. As sensitivity studies, nonparametric ANCOVA was performed for the studies, using the rank of the change from baseline of the primary outcome as the response, with treatment and site as factors and ranked baseline score as the covariate. Nonparametric ANCOVA test for Study 306 gave the Dunnett adjusted *p*-values of 0.005, 0.076 and 0.003 for dose groups of 100 mg, 200 mg, and 400 mg of DVS SR compared to placebo, respectively. Nonparametric ANCOVA test for Study 308 gave the Dunnett adjusted *p*-values of 0.003 and 0.014 for dose groups of 200 mg and 400 mg of DVS SR compared to placebo, respectively. Such analyses do not depend on model assumptions. Therefore, the normality assumption in the ANCOVA seemed to have very limited effect on the significance results.

For the primary endpoint, the results were supported by other pre-specified analyses in these studies (OC, MMRM, and ETRANK) at Week 8. According to Table 3.9, in Study 306, DVS SR dose groups of 100 mg, 200 mg, and 400 mg were all statistically significantly better than placebo in OC, MMRM and ETRANK analyses. In Study 308, both dose groups of DVS SR (200 mg and 400 mg) were statistically significantly better than placebo in the same analyses.

For the secondary endpoint, statistical significance was achieved for both 200 mg and 400 mg dose groups in Study 308. However, in Study 306, the procedure that the sponsor specified and used may not protect the overall family wise error rate for testing the primary and secondary endpoints. Thus, strictly speaking, statistical significance of the key secondary endpoint for 100 mg and 400 mg is difficult to conclude. However, the very small nominal *p*-values of the key secondary endpoint for these two doses seemed supportive.

3.2 Evaluation of Safety

See medical review for detail.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

4.1.1 Studies 309, 317, 320

The effect of sex on the treatment effect was explored by testing the significance of the treatment effect after the adjustment of sex alone, and sex by treatment interaction on the change from baseline of HAM-D₁₇. Table 4.1 gives the estimated treatment effect and nominal *p*-values in each gender group in Studies 309, 317 and 320. These *p*-values changed only slightly. Sex did not seem to have affected the significance level of the treatment on the primary endpoint.

Table 4.1 Exploratory Subgroups Analyses: Treatment Effect by Sex in Studies 309, 317 and 320 at Week 8 (LOCF Analysis)

Study Sex		Placebo	DVS SR	Ven ER	p-value ^a
Study 309					
Male	N	N=40	N=33	N=35	
	Effect size	-12.28	-13.95	-12.14	0.36, 0.93
Female	N	N=80	N=83	N=92	
	Effect size	-12.40	-13.16	-14.35	0.53, 0.095
Study 317					
Male	N	N=45	N=40	N=35	
	Effect size	-8.19	-11.15	-12.07	0.10, 0.038
Female	N	N=80	N=70	N=80	
	Effect size	-10.67	-10.08	-12.92	0.65, 0.07
Study 320					
Male	N	N=37	N=45		
	Effect size	-7.21	-9.48		0.15
Female	N	N=81	N=72		
	Effect size	-7.24	-8.91		0.13

a: This column gives one p-value in Study 320 and 2 p-values in Studies 309 and 317 for comparisons of DVS SR vs. placebo and Ven ER vs. placebo.

Source: FDA analysis.

The sample size was considerably larger in the female group. The above table shows that effect sizes were very similar between male and female, except in Study 317 in which males seemed to have larger effect than the females because females had larger placebo effect. All these results were derived after the adjustment of pooled center and baseline HAM-D₁₇ total score.

To consider the treatment effect in different ethnic groups, we note that in Study 309, there were 99% white patients, in Study 317 there were 82% white and in Study 320 there were 71% white. So the vast majority of the patients were white.

To consider the treatment effect in different age groups, we separated patients to below 65 (≤ 65) and above 65 (> 65) years of age. We note that in Study 309 there were 95% below 65, in Study 317 there were 99% below 65, and in Study 320 there were 99% below 65. So the vast majority of the patients were below 65 years of age.

4.1.2 Studies 306, 308

The effect of sex on the treatment effect was explored by testing the significance of the treatment effect after the adjustment of sex alone, and sex by treatment interaction on the change from baseline of the HAM-D₁₇ total score. Table 4.2 gives the treatment effect for each gender in Studies 306 and 308. These p-values changed only slightly. So sex did not seem to have affected the significance level of the treatment on the primary endpoint.

Table 4.2 Exploratory Subgroups Analyses: Treatment Effect by Sex in Studies 306 and 308 at Week 8 (LOCF Analysis)

Study Sex		Placebo	DVS SR 100 mg	DVS SR 200 mg	DVS SR 400 mg
Study 306					
Male	N	N=38	N=40	N=45	N=52
	Effect size	-7.84	-10.28 (p ^a = 0.30)	-7.95 (p ^a = 1.00)	-10.55 (p ^a = 0.19)
Female	N	N=80	N=74	N=71	N=61
	Effect size	-7.53	-10.72 (p=0.013)	-10.52 (p=0.022)	-11.04 (p=0.009)
Study 308					
Male	N	N=40		N=38	N=47
	Effect size	-9.17		-14.09 (p=0.026)	-11.52 (p=0.34)
Female	N	N=84		N=83	N=77
	Effect size	-9.53		-11.71 (p=0.15)	-12.49 (p=0.039)

a: this is the p-value of each dose group compared to the placebo, adjusted by Dunnett's method.

Source: FDA analysis.

The sample size was considerably large in the female group. The above table shows that the treatment effect sizes were comparable between males and females in Study 308, but females appeared to be more significant in Study 306. All these results were derived after the adjustment of pooled center and baseline HAM-D₁₇ total score.

To consider the treatment effect in different ethnic groups, we note that in Study 308 there were 99% white patients, in Study 306 there were 72% white. So the vast majority of the patients were white.

To consider the treatment effect in different age groups, we separated patients to below 65 (≤ 65) and above 65 (> 65) years of age. We note that in Study 308 there were 95% below 65, in Study 306 there were 99% below 65. So the vast majority of the patients were below 65 years of age.

4.2 Other Special/Subgroup Populations

Not available.

**Appears This Way
On Original**

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

5.1.1 Studies 309, 317, 320

These were flexible dose studies in which subjects were randomly assigned to receive DVS SR 200 to 400 mg, placebo, or a comparative treatment venlafaxine ER in US, Europe, South Africa and Taiwan to evaluate the efficacy and safety of DVS SR vs. placebo in adult outpatients with MDD. The primary efficacy analysis was performed using the ANCOVA on the change from baseline of the HAM-D₁₇ total score using LOCF data. The key secondary endpoint was the CGI-I score. The hierarchical testing procedure was used to control the overall Type I error rate.

The primary analysis using LOCF data in Studies 309, 317 and 320 did not support the claim of the effectiveness of DVS SR in the treatment of MDD. According my evaluation, the non-significance could be a result of unbalanced early dropout caused by adverse events. Removing these patients resulted in statistical significance of LOCF analyses in Studies 309 and 320. In addition, OC, MMRM and ETRANK analyses in these two studies also seemed to support the effectiveness of DVS SR in the reduction of the HAM-D₁₇ total score. However, due to unbalanced early dropout caused by adverse events, the risk benefit ratio should be carefully evaluated before giving a definite conclusion. The claim on the key secondary endpoint was not supported by the data in these studies.

5.1.2 Studies 306, 308

These were 8-week, fixed-dose studies with treatment arms DVS SR 100 mg, 200 mg, 400 mg and placebo in Study 306, DVS SR 200 mg, 400 mg and placebo in Study 308. The primary efficacy analyses on the change from baseline of the HAM-D₁₇ total score were performed using ANCOVA with LOCF data. Significance was adjusted by Dunnett's method. The key secondary efficacy variable was the CGI-I score.

The analysis results supported the efficacy claim of DVS SR in the treatment of MDD in both studies. In Study 306, DVS SR dose groups of 100 mg and 400 mg met the significance criterion while DVS SR 200 mg dose group did not. In study 308, both dose groups of DVS SR met the significance criterion. The significance results were supported by nonparametric methods as well as other pre-specified analyses (OC, MMRM, and ETRANK).

For the secondary endpoint, statistical significance was achieved for both 200 mg and 400 mg dose groups in Study 308. However, in Study 306, the procedure that the sponsor specified and used may not protect the overall family wise error rate for testing the primary and secondary endpoints. Thus, strictly speaking, statistical significance of the key secondary endpoint for 100 mg and 400 mg is difficult to conclude. However, the very small nominal p-values of the key secondary endpoint for these two doses seemed supportive.

5.2 Conclusions and Recommendations

The primary efficacy analysis using LOCF data in Studies 309, 317 and 320 did not support the claim of the effectiveness of DVS SR in the treatment of MDD. The non-significance could be a result of unbalanced early dropout caused by adverse events. Removing these patients resulted in significance of LOCF analyses in Studies 309 and 320. In addition, OC, MMRM and ETRANK analyses in these two studies also provided supportive information for the effectiveness of DVS SR in the treatment of MDD. However, due to unbalanced early dropout caused by adverse events, the risk benefit ratio should be carefully evaluated before giving a definite conclusion. The claim on the key secondary endpoint was not supported by the data in these Studies.

The analysis results in Studies 306 and 308 supported the efficacy claim of DVS SR in the treatment of MDD in both studies. In Study 306, for the primary endpoint, DVS SR dose groups of 100 mg and 400 mg met the significance criterion but the DVS SR 200 mg dose group did not. In Study 308, for the primary endpoint, both dose groups of DVS SR met the significance criterion. The significance results were also supported by nonparametric methods as well as other pre-specified analyses in these studies (OC, MMRM, and ETRANK). The key secondary endpoint in both dose groups of Study 308 met the statistical significance requirement. In Study 306, the data on the key secondary endpoint in dose groups of 100 mg and 400 mg seemed supportive.

Reference:

1. Dmitrienko, A., Offen, W. W., and Westfall, P. H. (2003). Gatekeeping strategies for clinical trials that do not require all primary effects to be significant. *Statist. Med.*, **22**: 2387-2400.
2. Chen, X., Luo, X., and Capizzi, T. (2005). The application of enhanced parallel gatekeeping strategies. *Statist. Med.*, **24**: 1385-97.

C:\Data\My Documents\NDA Review\NDA 2006\21992

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

/s/

Fanhui Kong
8/3/2006 03:43:55 PM
BIOMETRICS

Peiling Yang
8/3/2006 04:28:32 PM
BIOMETRICS

Kooros Mahjoob
8/3/2006 05:08:20 PM
BIOMETRICS

James Hung
8/3/2006 05:16:46 PM
BIOMETRICS