CENTER FOR DRUG EVALUATION AND RESEARCH

APPROVAL PACKAGE FOR:

APPLICATION NUMBER

21-392

Statistical Review(s)

STATISTICAL REVIEW AND EVALUATION

STATISTICAL KEY WORDS:	Dunnett's procedure, Dose-response, Decision rule, Co-primary endpoints
NDA NUMBER:	21-392
DATE RECEIVED BY CENTER:	August 22, 2001
DRUG NAME:	Cardizem — (Diltiazem Hydrochloride)
INDICATION:	Hypertension
SPONSOR:	Biovail Laboratories Inc.
DOCUMENTS REVIEWED:	NDA Volumes 1, 126-146 and electronic data sets
STATISTICAL REVIEWER:	John Lawrence, Ph.D. (HFD-710)
STATISTICAL TEAM LEADER:	Jim Hung, Ph.D. (HFD-710)
BIOMETRICS DIVISION DIRECTOR:	George Chi, Ph.D. (HFD-710)
CLINICAL REVIEWER:	Abraham Karkowsky, M.D. (HFD-110)
PROJECT MANAGER:	Colleen Locicero (HFD-110)

Table of Contents

List of Tables

1 Executive Summary of Statistical Findings

1.1 Conclusions and Recommendations

Based on the totality of evidence from one placebo-controlled clinical study, the 540 mg evening dose is effective in reducing trough diastolic blood pressure relative to placebo. There is some evidence that the 240 mg and 360 mg evening doses may be effective in reducing trough diastolic blood pressure relative to placebo. In addition, the 360 mg evening dose appears to be more effective than the 360 mg morning dose in reducing diastolic blood pressure measured between 6 am and 12 noon. All of the doses studied appear to be safe and well tolerated during the period studied.

1.2 Overview of Clinical Program and Studies Reviewed

There was one placebo-controlled study in support of the indication. In this study, a total of 478 patients were randomized and received at least one dose of study medication. Patients were independently randomized to one of 6 treatment groups. These groups were placebo, 120 mg PM, 240 mg PM, 360 mg AM, 360 mg PM, and 540 mg PM. In order to maintain the treatment blinding, all patients took three capsules in the morning and three capsules in the evening. However, some or all of the capsules contained no active treatment so that the dose for a specific patient would match the treatment to which they were randomized.

1.3 Principal Findings

Based on the evidence from one placebo-controlled clinical study, there appears to be strong evidence that the 540 mg evening dose is effective in reducing trough diastolic blood pressure relative to placebo ($p < 0.001$). Evening doses of 240 mg and 360 mg appeared to be effective relative to placebo ($p=0.034$ and 0.020 respectively). The strength of evidence is not as great with these lower doses and does not reach the level that would normally stand by itself for approval from a single study. The 120 mg evening dose was not shown to be significantly better than placebo at reducing trough blood pressure ($p=0.78$). Finally, the 360 mg evening dose appeared to be more effective at reducing diastolic blood pressure measured between 6 am and 12 noon compared to a morning dose of 360 mg ($p=0.0004$).

2 Statistical Review and Evaluation of Evidence

2.1 Introduction and Background

The investigators initiated this study with two primary objectives. The first goal was to compare each of the evening doses with placebo in the reduction of trough diastolic blood

pressure. The second objective was to compare the evening dose with a morning dose of the same amount (360 mg) in the reduction in DBP measured in the morning (6 am to 12 noon). All measurements in the window defining the measurement were obtained using an ABPM.

There were several secondary objectives. These included the comparison between treatment groups in changes in blood pressure or heart rate averaged over different time periods. For more details see Section 2.3.1 of this review.

2.2 Data Analyzed and Sources

The randomization was arranged so that 50% more patients would be randomized to each of the 360 mg groups (AM and PM) than would be randomized to each of the other four groups. Summaries of baseline demographics in each group and the number of patients that completed and withdrew for adverse events appear in Table 1. With the exception of the number of patients randomized to each group, there were no significant differences between any of the treatment groups with respect to any of these characteristics.

Table 1 Patient Disposition and Baseline Demographics (mean ± standard error)

	Placebo	120 PM	240 PM	360 AM	360 PM	540PM
Randomized	69	67	68	102	103	69
Completed	57	59	63	91	94	62
Withdrew for AE	3	3	3	2	1	4
Age	52 ± 10	52 ± 10	53 ± 9	54 ± 10	52 ± 8	51 ± 10
Male (%)	45 (65)	46 (69)	42 (62)	61 (60)	69 (67)	40 (58)
Race- Caucasian	45 (65)	35 (52)	42 (62)	66 (65)	67 (65)	47 (68)
Black	16 (23)	22 (33)	24 (35)	28 (28)	28 (27)	14 (20)
Weight (kg)	92 ± 18	87 ± 17	90 ± 20	94 ± 23	92 ± 18	92 ± 18
DBP (6PM-10PM)	99 ± 8	98 ± 10	100 ± 10	98 ± 9	96 ± 10	98 ± 10

[*Source: Tables 5 and 6 of Study Report*]

2.3 Statistical Evaluation of Evidence on Efficacy/ Safety

2.3.1 Sponsor's Results and Conclusions

The primary efficacy parameters were:

P1. Change from baseline to endpoint in trough DBP, as recorded by ABPM between 6PM and 10PM. Comparisons were made between each evening dose of study drug and placebo.

P2. Change from baseline to endpoint in mean DBP, as recorded by ABPM between 6AM and 12NOON. Comparisons were made between a morning dose of 360 mg and an evening dose of 360 mg of the study drug.

The summaries for the primary efficacy parameters appear in Table 2. The p-values corresponding to parameter P1 were adjusted for multiple comparisons using Dunnett's procedure. There were no adjustments to the overall type I error made for testing P1 and P2 simultaneously. See Section 2.3.2 of this review for more details of how these p-values were calculated.

Table 2 Primary efficacy parameters

Change in	Time period	Dose	LS Mean	Relative to	P-value
Trough DBP	6PM-10PM	Placebo	0.11		
Trough DBP	6PM-10PM	120PM	-1.92	Placebo	0.5180
Trough DBP	6PM-10PM	240PM	-4.26	Placebo	0.0208
Trough DBP	6PM-10PM	360PM	-4.38	Placebo	0.0081
Trough DBP	6PM-10PM	540PM	-8.02	Placebo	0.0001
DBP	6AM-12NOON	360AM	-6.27		
DBP	6AM-12NOON	360PM	-9.56	360AM	0.0004

[*Source: Tables 9, 10, 12, and 13 of Study Report. Not confirmed by FDA reviewer- see Section 2.3.4*]

There were eight secondary efficacy parameters. These were all changes from baseline in mean blood pressure or heart rate averaged over different time periods. These parameters all appear in Table 3, Table 4, and Table 5. For each parameter in a given time period, an adjustment for multiple comparisons was made for all pairwise comparisons relative to placebo using Dunnett's procedure. Trough PM is defined as 6PM-10PM for evening doses while trough AM is defined as 4AM-8AM for the morning dose.

Table 3 Secondary efficacy parameters- changes in mean DBP

Time period	Dose	Relative to	P-value
4AM-8AM	120PM	Placebo	0.0014
4AM-8AM	240PM	Placebo	0.0001
4AM-8AM	360AM	Placebo	0.0001
4AM-8AM	360PM	Placebo	0.0001
4AM-8AM	540PM	Placebo	0.0001
4AM-8AM	360PM	360AM	0.0022
Trough PM/Trough AM	360PM	360AM	0.0006
6AM-12NOON	120PM	Placebo	0.0014
6AM-12NOON	240PM	Placebo	0.0001
6AM-12NOON	360AM	Placebo	0.0001
6AM-12NOON	360PM	Placebo	0.0001
6AM-12NOON	540PM	Placebo	0.0001
6PM-10PM	360PM	360AM	0.0001
24-hour mean	120PM	Placebo	0.0052
24-hour mean	240PM	Placebo	0.0001
24-hour mean	360AM	Placebo	0.0001
24-hour mean	360PM	Placebo	0.0001
24-hour mean	540PM	Placebo	0.0001
24-hour mean	360PM	360AM	0.0862

[*Source: Table 15 of Study Report.*]

Table 4 Secondary efficacy parameters- changes in mean SBP

Time period	Dose	Relative to	P-value
6PM-10PM	120PM	Placebo	0.2667
6PM-10PM	240PM	Placebo	0.0075
6PM-10PM	360PM	Placebo	0.0121
6PM-10PM	540PM	Placebo	0.0001
6PM-10PM	360PM	360AM	0.0024
4AM-8AM	120PM	Placebo	0.0013
4AM-8AM	240PM	Placebo	0.0001
4AM-8AM	360AM	Placebo	0.0001
4AM-8AM	360PM	Placebo	0.0001
4AM-8AM	540PM	Placebo	0.0001
4AM-8AM	360PM	360AM	0.0008
Trough PM/Trough AM	360PM	360AM	0.0083
6AM-12NOON	120PM	Placebo	0.0008
6AM-12NOON	240PM	Placebo	0.0001
6AM-12NOON	360AM	Placebo	0.0001
6AM-12NOON	360PM	Placebo	0.0001
6AM-12NOON	540PM	Placebo	0.0001
6AM-12NOON	360PM	360AM	0.0004
24-hour mean	120PM	Placebo	0.0005
24-hour mean	240PM	Placebo	0.0001
24-hour mean	360AM	Placebo	0.0001
24-hour mean	360PM	Placebo	0.0001
24-hour mean	540PM	Placebo	0.0001
24-hour mean	360PM	360AM	0.6215

[*Source: Table 17 of Study Report.*]

Table 5 Secondary efficacy parameters- changes in mean HR

Time period	Dose	Relative to	P-value
6PM-10PM	120PM	Placebo	0.9998
6PM-10PM	240PM	Placebo	0.9999
6PM-10PM	360PM	Placebo	0.0503
6PM-10PM	540PM	Placebo	0.0517
6PM-10PM	360PM	360AM	0.1998
4AM-8AM	120PM	Placebo	0.9994
4AM-8AM	240PM	Placebo	0.7517
4AM-8AM	360AM	Placebo	0.1030
4AM-8AM	360PM	Placebo	0.1803
4AM-8AM	540PM	Placebo	0.0001
4AM-8AM	360PM	360AM	0.7788
Trough PM/Trough AM	360PM	360AM	0.0238
6AM-12NOON	120PM	Placebo	0.4277
6AM-12NOON	240PM	Placebo	0.0186
6AM-12NOON	360AM	Placebo	0.0038
6AM-12NOON	360PM	Placebo	0.0001
6AM-12NOON	540PM	Placebo	0.0001
6AM-12NOON	360PM	360AM	0.1365
24-hour mean	120PM	Placebo	0.9683
24-hour mean	240PM	Placebo	0.1469
24-hour mean	360AM	Placebo	0.0001
24-hour mean	360PM	Placebo	0.0001
24-hour mean	540PM	Placebo	0.0001
24-hour mean	360PM	360AM	0.3478

[*Source: Table 19 of Study Report.*]

From Table 3, it is clear that over each of these time periods, each dose of study drug beat placebo on a pairwise comparison basis. The 360AM dose was significantly better than the 360PM dose in the trough PM/trough AM comparison and in the 6PM-10PM comparison and numerically better in the 24-hour mean comparison. On the other hand, the 360PM dose was better than the 360AM dose during the 4AM-8AM time period. Qualitatively similar results were found for SBP in Table 4. The effects on HR, shown in Table 5, were less dramatic. Significant changes in HR were only observed with the higher doses and only during certain time periods.

No patient died within the study or within 30 days of the last dose of study drug. A total of 12 patients experienced SAEs during the study: 8 during the placebo run-in period and 4 during the double-blind period. None of these were considered to be related to study drug. 16 patients withdrew during the double-blind study period for adverse events. Of these, 11 were considered to be related to study medication. The most common treatment-emergent AEs were

headache (11.7%), upper respiratory tract infection (5.6%), and lower limb edema (5.4%). The incidence rates did not appear to be dose related.

2.3.2 Statistical Methodologies

All efficacy analyses presented in this review were done on the ITT population. For patients with missing values, the last observation was carried forward. However, ABPM measurements were only taken at baseline and at the end of the study. This had the effect of eliminating all patients who did not complete the study from the analysis of the primary efficacy variable. In the evening doses and the placebo group, 335 patients completed the study and were included in the primary analysis out of 376 patients randomized. Table 6 summarizes the reasons that patients dropped out by treatment group. Both the number of patients who dropped out prematurely and the distribution for the reasons listed in the table appears to be roughly the same in each group.

Table 6 Reasons for premature withdrawal from double-blind period

	Placebo	120 PM	240 PM	360 AM	360 PM	540PM
Randomized	69	67	68	102	103	69
Completed	57	59	63	91	94	62
Withdrew for AE	3	3	3	2	1	4
Noncompliance	1	1	0	3	4	1
Consent withdrawn	3	0	0	4	1	0
Lack of efficacy	3	1	1	1	0	0
Other	2	3	1	1	3	2

[*Source: Tables 5 of Study Report*]

Separate ANCOVA models were fit for each efficacy parameter. The ANCOVA model included treatment group and study site as main effects and baseline value as a covariate. The treatment-by-baseline interaction and treatment-by-site interactions were examined, but removed from the model if not significant at level 0.10. The final model included only baseline, treatment group, and study site since none of the interactions were found to be significant.

For a specific time period, multiple comparisons with the placebo were performed using Dunnett's procedure. No adjustment was made for simultaneously testing multiple hypotheses across different time windows and across different efficacy parameters.

2.3.3 Detailed Review of Individual Studies

Male or female patients between the ages of 18 and 70 with moderate to severe hypertension were enrolled. Patients entered a 3 to 4 week single-blind, placebo run-in period. Afterwards, qualifying patients entered a 7-week double-blind treatment phase. A 36-hour

ABPM was to be performed at the end of the placebo run-in period (baseline) and repeated at the end of the double-blind treatment period.

Some concomitant medications were excluded during the study period including anti-hypertensive drugs and NSAIDs.

2.3.4 Statistical Reviewer's Findings

The results for the primary efficacy parameters found by the reviewer were slightly different than, but qualitatively similar to the sponsor's results presented in this review. There is a technical issue regarding the validity of the use of Dunnett's procedure. A condition should be checked regarding the covariance matrix of the differences to verify the validity of Dunnett's procedure[1]. For this data, the condition was not satisfied. This reviewer did the analysis two different ways: the first way ignores this issue and blindly applies the formula for Dunnett's procedure, the second way is to use a simulation based method to find the critical values[2]. The results for the primary efficacy parameter P1 (change in trough DBP during 6PM-10PM) appear in Table 7. There does not seem to be any practical difference in the p-values resulting from these two methods in this case.

Table 7 Analysis of primary efficacy parameter P1 (change in trough DBP)

Dose	LS Mean	P-value relative to Placebo: Dunnett's	P-value relative to Placebo: Simulation-based
Placebo	0.06		
120PM	-1.40	0.78	0.78
240PM	-4.13	0.035	0.034
360PM	-4.10	0.020	0.020
540PM	-7.27	< 0.0001	< 0.0001

[*Source: FDA analysis*]

The p-values in Table 7 differ from the sponsor's because of the way that study site was included as a covariate. In the FDA analysis, no pooling of centers was done. On the other hand, the sponsor's analysis pooled centers enrolling fewer than 12 patients to artificially create two larger pseudo-centers. All study sites were in the United States, which precluded pooling of centers base on country. The FDA analysis found the treatment-by-center interaction was significant at level 0.13, but not at the pre-specified level for including the term (0.1). Hence, the FDA analysis did not include the interaction term. The sponsor found a significant treatment-by-center interaction, but nonetheless did not include the term in the primary analysis because no discernable pattern could be detected for the interaction [*p. 60 of Study Report*].

[1] Dunnett, C. W. (1964). New tables for multiple comparisons with a control. Biometrics, 20, 482-491

[2] Edwards, D. and Berry, J. J. (1987). The efficiency of simulation-based multiple comparisons. Biometrics, 43, 913-928.

Table 8 shows the results for the FDA analysis of the second primary efficacy variable (change in DBP during 6AM-12NOON) comparing the evening dose of 360 mg to the morning dose of the same strength. As in the analysis of P1, the sponsor pooled small study sites while the FDA analysis does not. This explains the small numerical differences between the FDA's and the sponsor's analyses.

Table 8 Analysis of primary efficacy parameter P2 (change in DBP during 6AM-12NOON)

Dose	LS Mean	P-value relative to 360 mg AM
360 mg AM	-6.18	
360 PM	-9.68	0.0003

[*Source: FDA analysis*]

2.4 Findings in Special/Subgroups Populations

The primary analysis on change from baseline in trough DBP was investigated in four subgroups. Two separate ANCOVA models that included the terms in the model for the primary analysis as well as additional terms for each covariate and the covariate-by-treatment interaction were used to estimate the effects in these subgroups. The results appear in Table 6. Females appeared to respond to the drug better than males and this was true generally across doses. Similarly, patients who were 55 or older responded better than patients under 55 years old responded.

Table 9 Change in trough DBP (6PM-10PM) in subgroup populations.

Subgroup	Dose	LS Mean	SE
Age <55	Placebo	1.02	1.52
	120PM	-0.83	1.70
	240PM	-3.76	1.48
	360PM	-2.41	1.24
	540PM	-4.53	1.60
Age ≥55	Placebo	-1.44	1.95
	120PM	-1.74	1.70
	240PM	-4.70	1.89
	360PM	-6.64	1.54
	540PM	-10.44	1.74
Male	Placebo	0.87	1.54
	120PM	-1.32	1.50
	240PM	-3.23	1.52
	360PM	-4.44	1.22
	540PM	-6.06	1.59
Female	Placebo	-1.53	2.02
	120PM	-1.84	2.22
	240PM	-5.73	1.94
	360PM	-3.32	1.72
	540PM	-9.06	1.89

[*Source: FDA analysis*]

2.5 Statistical and Technical Issues

Dunnett's procedure was used to adjust for multiple comparisons within a single time period and parameter, but no adjustment was made for testing across different parameters or time periods. If one believes that the total experiment-wise error should be controlled across all secondary endpoints, or indeed across all primary and secondary endpoints simultaneously, then the correctly adjusted p-values would be much higher than those presented here. Because this is such a delicate issue and needs to be pre-specified to be done correctly, it is necessary to either ignore this issue entirely or simply interpret all of the nominal p-values (for the secondary endpoints) as descriptive statistics. Although two primary hypotheses were pre-specified, only one set of these hypotheses appears to be relevant for approval (namely the comparisons of the evening doses with placebo). Hence, one can argue that for approval, the error should be controlled at 5% simultaneously only on the four comparisons with respect to parameter P1 and this adjustment was done correctly.

2.6 Statistical Evaluation of Collective Evidence

Unless there is substantial external evidence, the agency normally sets a very high standard for approval of a drug based on a single study. The 540 mg dose probably meets that standard with a p-value less than 0.0001. The point estimate of the placebo-corrected change from baseline in trough DBP for the 240 mg and 360 mg PM doses were approximately 4 mm Hg, which seems to be generally accepted as a clinically meaningful reduction. However, the corresponding p-values failed to reach the level required for approval for a single study (0.034 and 0.020 respectively).

The 360 PM dose was significantly better than the 360 mg AM dose at reducing DBP during 6AM to 12NOON. However, this was at the expense of being significantly worse than the AM dose during other time periods (see Table 3 of this review and discussion following the table). Although the former time period was pre-specified as the time period for the primary analysis comparing these two regimens, the medical division will have to make a clinical judgement as to the clinical relevance of these results.

2.7 Conclusions and Recommendations

All doses of the drug that were studied appear to be safe after 7 weeks of treatment for reduction of hypertension. The level of evidence required to establish efficacy in a single study was reached only for the highest dose studied (540 mg PM).

2.8 Appendix of Individual Studies Reviewed

Not applicable.

2.9 Appendix of Statistical and Technical Discussions

Not applicable.

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

/s/

John Lawrence
2/13/02 04:03:07 PM
BIOMETRICS

James Hung
2/13/02 04:22:26 PM
BIOMETRICS

George Chi
2/13/02 04:34:52 PM
BIOMETRICS