CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

21-813

STATISTICAL REVIEW(S)

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

Statistical Review and Evaluation

CLINICAL STUDIES

NDA/Serial Number:	21-813
Drug Name:	Elestrin™ (Estradiol Transdermal Gel)
Indication(s):	Treatment of moderate-to-severe vasomotor symptoms.
Applicant:	BioSante Pharmaceuticals, Inc.
Date (s):	Submitted: 2/16/2006
	PDUFA: 12/16/2006
Review Priority:	Standard
Biometrics Division:	Division of Biometrics III (HFD-725)
Statistical Reviewer:	Mahboob Sobhan, Ph.D. (HFD-725)
Concurring Reviewer:	Lisa Kammerman, Ph.D. (HFD-725)
Medical Division:	Division of Reproductive and Urological Drug Products (HFD-580)
Clinical Team:	Shelley Slaughter, M.D. (HFD-580)
	Theresa H. Van der Vlugt, M.D. (HFD-580)
Project Manager:	George Lyght (HFD-580)
Keywords:	NDA review, Clinical studies, subgroup analyses, treatment by baseline interaction, multiple comparison.

Table of Contents

1.0 EXECUTIVE SUMMARY

1.1 Conclusion and Recommendations

Based on the efficacy data submitted from Phase 3 study EST005, our analysis showed a statistically significant reduction in daily moderate-to-severe hot flush rate and severity for all doses of Elestrin starting at week 5 and maintained through week 12.

[redacted]

From a statistical perspective, this application provided adequate data to support Elestrin in the treatment of hot flush frequency and severity [redacted] in postmenopausal women.

1.2 Brief Overview of Clinical Studies

The applicant, BioSante Pharmaceuticals Inc., reports clinical data to support Elestrin for [redacted] indications: Vasomotor symptoms (VMS) [redacted]. Data from two studies: EST004, a Phase 2 study and EST005, a Phase 3 study were submitted to demonstrate the safety and effectiveness of Elestrin. Study EST004 was not adequately powered to demonstrate statistical significance and therefore not considered a pivotal study to support efficacy. Study EST005 was a parallel-group, placebo-controlled, randomized study designed to detect a clinically meaningful difference of ≥2.0 between Elestrin dose groups and placebo. The protocol called for randomizing 127 subjects per group equally in the following groups: Elestrin 0.87 g/day, Elestrin 1.7 g/day, Elestrin 2.6 g/day and placebo, but only 69 subjects were randomized in the Elestrin 2.6 g/day dose group. This resulted in an unbalanced number of subjects among treatments. The explanation was that the effect size for the higher dose would be 3.0 instead of 2.0, and therefore, 69 subjects would provide adequate power to test the null hypothesis of no difference for this group compared to placebo.

The efficacy outcomes were the evaluation of hot flushes (rated as mild, moderate, and severe) and severity (rated as 1, 2, and 3) using a daily diary. The primary efficacy endpoints were the mean changes in the hot flush rate and severity from baseline to weeks 4 and 12. An additional efficacy outcome, vulvar vaginal atrophy symptoms (VVA) was also evaluated as measured by three components: rate of most bothersome moderate to severe symptoms based on five questions (dryness, irritation, pain passing urine, pain with sexual activity, and bleeding with sexual activity rated as none=0, mild=1, moderate=2, and severe=3) using vaginal self-health assessment questionnaire, vaginal PH and vaginal maturation index.

The study was designed to test the superiority hypothesis with respect to two co-primary endpoints:

hot flush rate and severity. It was not planned to test the difference between Elestrin and placebo in the most bothersome moderate-to-severe atrophy symptoms as a third co-primary endpoint. The protocol specified statistical analysis methods included analysis of covariance (ANCOVA) to evaluate pair-wise comparisons with adjustment for multiple dose comparisons using two analysis populations: intent-to-treat with imputation for missing post-baseline diaries and without imputation (observed data at weeks 4 and 12).

1.3 Statistical Issues and Principal Findings

Our review focused on several statistical issues: the impact of missing post-baseline diary data, adjustment for multiple comparisons (pair-wise comparison of each dose group versus placebo) multiplicity (multiple endpoints), and adequacy of the study power with regards to all primary endpoints. Missing diaries were reported in less than 7% of the subjects (ranging from 3% to 7% across treatment groups) and did not appear to follow any missing pattern, i.e., missing either due to adverse events or lack of efficacy. The efficacy results using last-observation-carried-forward approach (LOCF) and observed (completers at endpoint) analysis population were similar. The adjustment for multiple dose group comparisons were made for the evaluation of hot flush rates, severity (VMS), and most bothersome atrophy symptoms (VVA), but no adjustment for multiplicity (multiple components) was made for the VVA.

Considering all the above statistical issues and based on the applicant's data and our independent analysis (adjusting for multiple comparisons), the efficacy results could be summarized as follows:

(1) For the VMS indication, the two highest doses of Elestrin: 1.7 g/day and 2.6 g/day, respectively, demonstrated both clinically meaningful and statistically significant ($p<.001$) reductions in moderate-to-severe hot flush rate and severity at week 4 and maintained through week 12, compared to placebo. Elestrin 0.87 g/day dose, however, demonstrated only marginally significant ($p=.0511$) reductions in both rate and severity at week 4.

(2) At week 5, all three doses of Elestrin demonstrated clinically and statistically significant ($p<.001$) reduction in host flush rate and severity, and maintained through week 12.

2.0 INTRODUCTION

2.1 Overview

The applicant, BioSante Pharmaceuticals, Inc., seeks approval for Elestrin, a topical estradiol in a gel, in the treatment of moderate to severe vasomotor symptoms (VMS) [ ] associated with menopause. BioSante claims that this formulation is applied directly to the skin and will have potential advantages over oral or other transdermal patches formulations in terms of reduction in skin reactions observed with transdermal patch formulations.

To support the safety and efficacy of Elestrin, data from two studies: a Phase 2 study (EST004) and a Phase 3 study (EST005) were submitted in this application. At the end of the Phase 2 meeting, the Division determined that study EST004 was not adequately powered to determine the lowest effective dose of Elestrin, and recommended a larger study. In response, the sponsor conducted study EST005 to find the lowest effective dose as well as safety and efficacy of Elestrin in a much larger patient population as shown in Table 2.1. This review will evaluate the efficacy data from study EST005 only. Pertinent safety data from study EST004 will be included in the clinical reviewer's report.

Table 2.1
Summary of Clinical Studies

Study#	Objectives	Study Design	Study Regimen	Number Randomized
EST004	Safety and efficacy	Multi-center, Double-blind, Placebo-controlled, Phase 2.	Elestrin 0.625 g/day Elestrin 1.25 g/day Elestrin 2.50 g/day Placebo	41 39 38 42
EST005	Safety and Efficacy	Multi-center, Double-blind, Placebo-controlled, Parallel-group, Phase 3	Elestrin 0.87 g/day Elestrin 1.7 g/day Elestrin 2.60 g/day Placebo	136 142 69 137

2.2 Indication

Elestrin 0.87 g/day and Elestrin 1.7 g/day are indicated in the:

1. *Treatment of moderate to severe vasomotor symptoms associated with menopause.*

[ ]

3.0 STATISTICAL EVALUATION

3.1 Overview of Study EST005

3.1.1 Study Design

Study EST005 was a multi-center, randomized, placebo-controlled, parallel-group study conducted to demonstrate safety and efficacy of Elestrin (0.87 gm, 1.7 g, and 2.6 g) compared to placebo in the treatment of vasomotor symptoms (VMS) and vulvar and vaginal atrophy (VVA) symptoms in postmenopausal women. Initially, the study was designed to include two doses of Elestrin (1.7 g and 2.6 g) with equal numbers of subjects, but the design was modified to include a lower dose after the Division's recommendation that an ineffective lower dose would not be demonstrated if both higher doses demonstrate effectiveness. By an amendment (7), the protocol reduced the number of subjects in the higher dose by half without reasonable explanation and added the lower dose group with the same number as placebo, resulting in an unbalanced number of treated subjects.

3.1.2 Efficacy Outcome

The efficacy measurements included the followings:

1) Vasomotor Symptoms: Vasomotor symptoms were evaluated by hot flush frequency and severity using a recorded daily diary as follows:

Hot Flushes	Severity Score	Classification
Mild	1	Sensation of heat without perspiration.
Moderate	2	Sensation of heat with perspiration, and able to continue activity.
Severe	3	Sensation of heat and sweating, causing the subject to stop an activity until the event passed.

Hot flush rate was calculated as the total number of moderate to severe hot flushes divided by the number of those 7 days with diary entries completed.

Hot flush severity was calculated as the sum of the average daily hot flush severity ratings (using severity score above) divided by the number of those 7 days with diary entries completed.
The primary efficacy endpoints were the mean change from baseline to weeks 4 and 12 in moderate to severe hot flush rate and severity.

2) Vulvar and Vaginal Atrophy (VVA) Symptoms: VVA was evaluated by vaginal PH, vaginal maturation index (calculated from the percentages of parabasal, intermediate, and superficial cells), and most bothersome moderate to severe symptoms based on four questions (dryness, irritation, pain passing urine, pain with sexual activity, and bleeding with sexual activity) rated as none (0), mild (1), moderate (2), and severe (3) from the vaginal health self-assessment questionnaire.

As per the Division's guidance for VVA[ ], the most bothersome moderate-to-severe vulvovaginal atrophy symptoms were considered primary among the three components stated above and both PH and MI should be limited to subjects who meet the following criteria: 1) subjects who had at least one moderate-to-severe symptom of vulvar and vaginal atrophy that was most bothersome to her, 2) vaginal PH>5.0, and 3) had superficial cells ≤5% on a vaginal smear.

Secondary Efficacy: The secondary efficacy variables included a responder analysis of both co-primary endpoints defined as proportion of subjects with ≥50% to 100% reductions from baseline to weeks 4 and 12, global assessment of efficacy, subject opinion survey, change from baseline in UQOL total score, MENQOL quality of life domain to week 12.

3.1.3 Determination of Sample Size

Using a clinically meaningful difference of ≥2.0 in daily hot flush rate between Elestrin and placebo in the mean change from baseline to week 4 with a standard deviation of 5.0, the protocol called for a planned sample size of 127 for the lower two dose groups of Elestrin and placebo, with 80% power at alpha=0.05.

3.1.4 Statistical Methods

For comparison of treatment groups with respect to both co-primary endpoints, the statistical methods included ANOVA models including center, treatment, and center-by-treatment interaction as factors. Pair-wise comparisons were to be reported as least square (LS) means, adjusting for multiple comparison by Dunnett's test.

3.1.5 Reviewer's Comments on the Design

Study EST005 randomized adequate number of subjects to test the superiority hypothesis with respect to two VMS co-primary endpoints in the two lowest dose groups of Elestrin. [ ]

[ ]

3.2 Study Results

3.2.1 Subject Disposition

A total of 484 subjects were randomized into the double-blind treatment period as follows: Placebo (137), Elestrin 0.87 g (136), Elestrin 1.7 g (142), and Elestrin 2.6g (69). Of these, 27 (5.6%) subjects discontinued prematurely, with the reasons for discontinuation being adverse events (9), withdrawn consent (5), lack of efficacy (3), non-compliance (2), lost to follow-up (2), estradiol >2 ng/L (2), and other non-specified reasons (4). For the evaluation of the VMS indication, the ITT analysis population included adequate number of subjects in the lowest doses of Elestrin and placebo group, while for the [ ]

Table 3.2.1 Disposition of Subjects

Subjects	Treatment groups				
	Placebo	Elestrin 0.87 g	Elestrin 1.7 g	Elestrin 2.6 g	Total
Received study drug	137	136	142	69	484
Completed study	128	132	133	64	457
Discontinued	9	4	9	5	27
Data set for efficacy analysis(ITT)*:					
VMS	137	136	142	69	484
[redacted]					
* ITT population included all randomized subjects who received treatments and had diary response for at least 1 full day					

3.2.2 Patient demographics and Baseline characteristics

The baseline characteristics of patients randomized in this study were comparable between treatment groups with regards to age, race, BMI, height, weight, and medical history. There were no significant between treatment differences with respect to menopausal history or prior hormone therapy either.

3.2.3 Efficacy

Hot Flush Rate: The study was designed to evaluate two co-primary endpoints: Mean change from baseline in 1) daily moderate-to-severe hot flush rates, and 2) daily hot flush severity to week 4 and week 12. In addition, the sponsor also evaluated vulvovaginal atrophy symptoms as another primary endpoint based on subgroup analysis, although sample size calculation was based on the rates and severity endpoints only.

We performed a statistical analysis similar to the sponsor's analysis using analysis of covariance (ANCOVA) with factors for baseline, treatment, site, and baseline by treatment interaction. For treatment comparisons with placebo, Dunnett's test was used for controlling type-I error. Our analysis was also based on ITT population using last observation carried forward (LOCF) for missing post-baseline data. We used LOCF because the percentages of subjects with post baseline missing diary ranged from 3%-7%. It was similar across treatment groups and did not appear to follow any systematic pattern that could either be considered as missing at random or otherwise.

Results of our analyses are shown in Table 3.2.3. At week 4, all three treatment groups showed a reduction in moderate-to-severe hot flush rates, but the change was statistically significantly (<.001) greater only in the two higher dose groups: Elestrin 1.7 g/day and Elestrin 2.6 g/day, compared to placebo. For subjects receiving Elestrin 0.87 g/day, the change was marginally statistically significant

(p=.0511) at week 4. However, beginning at week 5 both clinically (mean change of >2) and statistically meaningful differences were seen for all doses of Elestrin compared to placebo, and were maintained through week 12. It appeared 0.87 g/day is the lowest effective dose of Elestrin for treating moderate-to-severe hot flushes, since there were no differences between Elestrin doses at week 4 and 12.

Results using evaluable ITT population using completers (not shown in Table 3.2.3), were similar to ITT using LOCF analysis population.

Hot Flush Severity: For hot flush severity, our analysis also demonstrated similar results. At week 4, no statistically significant reduction in severity was noted for Elestrin 0.87 g/day compared to placebo. Starting at week 5, however, all three doses of Elestrin were both clinically (>0.72 effect size) and statistically (p<.001) superior to placebo in reducing severity of hot flush severity, and the reduction was maintained through week 12.

Vulvar and Vaginal Atrophy (VVA) Symptoms: [ ], as per Division's guidance document, three endpoints: the most bothersome moderate-to-severe atrophy symptoms (composite), vaginal PH and vaginal maturation index (MI) were evaluated in subjects who met three HT guidance criteria, i.e., in subjects who had at least 1 moderate-to-severe symptom of vulvar and vaginal atrophy which the subject self-identified as most bothersome to her, vaginal PH>5 and ≤5% superficial cells on a vaginal smear.

Table 3.2.3
Change from Baseline to Week 12 in the Mean Number of Moderate and Severe Hot Flushes: ITT-LOCF Population, Study EST005

Efficacy	Treatment Groups			
	Placebo (N=137)	Elestrin 0.87 g (N=136)	Elestrin 1.7 g (N=142)	Elestrin 2.6 g (N=69)
Moderate to Severe Hot Flushes				
Baseline Mean (±SD)	**13.47 ± 4.5**	**13.30 ± 4.6**	**13.10 ± 6.5**	**12.87 ± 4.4**
Weekly Mean (LS Mean change)+				
1	9.23 (-3.98)	9.63 (-3.45)	10.59(-2.38)**	9.94 (-2.92)
2	8.57 (-4.53)	8.43 (-4.58)	8.26 (-4.60)	7.17 (-5.68)
3	8.14 (-4.95)	7.47 (-5.61)	6.16 (-6.75)**	4.75 (-8.30)***
4	7.91 (-5.14)	6.55 (-6.50)	4.87 (-8.00)***	3.69 (-9.32)***
5	7.83 (-5.14)	5.50 (-7.47)***	4.03 (-8.81)***	3.19 (-9.83)***
6	7.60 (-5.37)	5.20 (-7.70)***	3.54 (-9.30)***	2.80 (-10.19)***
7	7.31 (-5.75)	4.67 (-8.25)***	3.21 (-9.69)***	2.30 (-10.79)***
8	7.32 (-5.67)	4.55 (-8.33)***	3.04 (-9.79)***	2.25 (-10.77)***
9	7.40 (-5.61)	4.46 (-8.43)***	2.82 (-10.04)***	1.90 (-11.14)***
10	7.40 (-5.57)	4.22 (-8.61)***	2.68 (-10.16)***	2.00 (-11.04)***
11	7.23 (-5.67)	4.15 (-8.63)***	2.61 (-10.16)***	2.04 (-10.96)***
12	7.30 (-5.35)	4.00 (-8.50)***	2.50 (-10.02)***	2.05 (-10.66)***
Hot Flush Severity				
Baseline Mean (±SD)	**2.41±0.32**	**2.42±0.32**	**2.40±0.27**	**2.41±0.32**
Weekly Mean (LS Mean change)+				
1	2.26 (-0.16)	2.27 (-0.15)	2.31 (-0.09)	2.32 (-0.10)
2	2.20 (-0.19)	2.18 (-0.22)	2.15 (-0.25)	2.02 (-0.40)*
3	2.14 (-0.25)	2.08 (-0.34)	1.88 (-0.52)**	1.71 (-0.71)***
4	2.12 (-0.24)	1.93 (-0.45)	1.70 (-0.67)***	1.45 (-0.96)***
5	2.12 (-0.22)	1.85 (-0.50)**	1.59 (-0.77)***	1.34 (-1.05)***
6	2.07 (-0.27)	1.77 (-0.57)**	1.46 (-0.90)***	1.26 (-1.14)***
7	2.04 (-0.30)	1.70 (-0.64)**	1.34 (-1.01)***	1.10 (-1.31)***
8	2.06 (-0.27)	1.65 (-0.67)**	1.27 (-1.07)***	1.02 (-1.39)***
9	2.06 (-0.27)	1.64 (-0.67)**	1.28 (-1.08)***	0.93 (-1.47)***
10	2.10 (-0.22)	1.60 (-0.70)**	1.18 (-1.16)***	0.87 (-1.52)***
11	2.06 (-0.25)	1.55 (-0.76)***	1.19 (-1.14)***	0.92 (-1.47)***
12	2.05 (-0.26)	1.52 (-0.77)***	1.15 (-1.17)***	0.86 (-1.51)***

+ Change from baseline; *p<.05, **p<.01, *** P<.0001, for pair wise comparison (adjusting for multiple comparison using Dunnett's test) with placebo for the differences in LS means from ANCOVA model with factors for baseline, treatment, center, and baseline by treatment interaction (significant for change in hot flush rate only).

3.2.4 Reviewer's Comments on the Efficacy Results

Adjusting for multiple dose comparisons, our analysis confirmed the sponsor's efficacy results of a statistically significant reduction in hot flush rate and severity for all doses of Elestrin, starting as early as week 5 and maintained through week 12.

4.0 SUMMARY AND CONCLUSIONS

This review evaluated the efficacy data from study EST005 in support of Elestrin compared to placebo for the treatment of moderate-to-severe hot flushes in postmenopausal women. Study EST005 was a randomized, placebo-controlled, parallel-group, dose-ranging study. The study was planned to randomize 127 per group assuming an effect size of ≥2.0 (clinically meaningful difference between Elestrin and placebo) with 80% power, but only 69 subjects were randomized in the highest dose group of Elestrin. With only 69 subjects, the power was reduced to approximately to 50%.

Based on the submitted data, our analysis showed that all doses of Elestrin 0.87 g/day, 1.7 g/day, and 2.6 g/day were statistically significant ($p<.001$) in reducing moderate-to-severe hot flush rate and severity, compared to placebo starting at week 5 and maintained through week 12.

From a statistical perspective, the efficacy data provided in this application do support the VMS indication.

Redacted 3 page(s)

of trade secret and/or

confidential commercial

information from

Statistical Review

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

/s/

Mahboob Sobhan
12/12/2006 11:09:43 AM
BIOMETRICS

Lisa A. Kammerman
12/12/2006 11:12:54 AM
BIOMETRICS
I concur with Dr. Sobhan's review.