CENTER FOR DRUG EVALUATION AND RESEARCH

APPROVAL PACKAGE FOR:

APPLICATION NUMBER

21-371

Statistical Review(s)

Screening of New NDA
Division of Biometrics II

Date: July 10, 2001

NDA #: 21-371

Priority Classification: 3S

Trade Name: Estrasorb

Applicant: Novavax, Inc.

Generic Name: estradiol

Date of Submission: 6/29/01

Indication:

No. of Controlled Studies: 1

User Fee Goal Date: 6/29/02

Date of 45-Day Meeting: 8/8/01

Medical Officer: Theresa van der Vlugt, M.D. (HFD-580)

Project Manager: Diane Moore (HFD-580)

Screened by: Kate Meaker, M.S.

Volume numbers in statistical section: 1.001, 1.025 – 1.051, 1.053

Anticipated Review Completion Date: 3/31/02

Comments:

1. Vol. 1.054 contains a CD-ROM that was sent to the electronic document room. This contains the SAS transport files and descriptions needed for my review.

2. In Study E99-1, at Week 4 there is a potential interaction between strata (uterus intact? Yes/No) and treatment effect (p-value = 0.073). I will discuss this in my review.

3. This application is fileable.

CHECKLIST

Item	Check (NA if not applicable)
Index sufficient to locate necessary reports, tables, etc.	Yes
Original protocols & subsequent amendments available in the NDA	Yes
Designs utilized appropriate for the indications requested	Yes
Endpoints and methods of analysis spelled out in the protocols	Yes
Interim analyses (if present) planned in the protocol and appropriate adjustments in significance level made	NA
Appropriate references included for novel statistical methodology (if present)	NA
Sufficient data listings and intermediate analysis tables to permit statistical review	Yes
Data from primary studies on diskettes and/or CANDA submitted	Yes - EDR
Intent-to-treat analysis	Yes
Effects of dropouts on primary analyses investigated	Yes
Safety and efficacy for gender, racial, and geriatric subgroups investigated	No

BRIEF SUMMARY OF CONTROLLED CLINICAL TRIAL

Study Number (Dates Conducted)	Number of Centers (Locations)	Total Sample Size	Design	Duration of Treatment
E99-1 (10/99 – 2/01)	21 (U.S.)	Estrasorb 7.5 mg/day (n=100) Placebo (n=100)	Multicenter, Double-blind, Randomized, Placebo-controlled, Parallel arm	12 Weeks

---

Statistical Reviewer

Concur: Dr. Welch

cc:
Archival NDA #21-371
HFD-580
HFD-580/TvanderVlugt, DMoore, SAllen
HFD-715/ENevius, MWelch, KMeaker

**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/
- - - - - - - - - - - - - - - - - - - -
Katherine Meaker
7/11/01 09:49:22 AM
BIOMETRICS

Mike Welch
7/16/01 09:06:22 AM
BIOMETRICS

APPEARS THIS WAY
ON ORIGINAL

DEPARTMENT OF HEALTH AND HUMAN SERVICES
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
OFFICE OF BIOSTATISTICS

Statistical Review and Evaluation

ADDENDUM TO 4/22/02 REVIEW

CLINICAL STUDIES

NDA:	21-371
Name of drug:	Estrasorb
Applicant:	Novavax, Inc.
Indication:	women
Documents reviewed:	Vols. 1.001, 1.025-1.051, 1.053;
	\\CDSESUB1\N21371\N_000\2001-06-29
Project manager:	George Lyght (HFD-580)
Clinical reviewer:	Theresa van der Vlugt, M.D. (HFD-580)
Dates:	Received 9/12/02; User fee (10 month) 7/12/02; HFD-580 requested date 6/15/02
Statistical reviewer:	Kate Meaker, M.S. (HFD-715)
Statistics team leader:	Mike Welch, Ph.D. (HFD-715)
Biometrics division director:	Ed Nevius, Ph.D. (HFD-715)

Keywords: NDA review, clinical studies, one study application

**APPEARS THIS WAY
ON ORIGINAL**

1 EXECUTIVE SUMMARY OF STATISTICAL FINDINGS

1.1 CONCLUSIONS AND RECOMMENDATIONS

My initial review of the clinical study to support approval of Estrasorb was completed on 4/22/02. [   ] It was resubmitted on 9/12/02, with no new clinical data. This addendum to my review was requested by the Medical Officer to confirm Estrasorb was statistically significantly better than placebo using a symptom severity score endpoint. The efficacy evaluation in the initial review covered symptom frequency and severity index endpoints.

The severity index was calculated a s a weighted sum of vasomotor symptoms, with mild receiving a weight of 1, moderate a weight of 2, and severe a weight of 3. The index did not divide by the total number of vasomotor symptoms experienced by a patient. The only difference in the calculation of the severity score in this review is that it does divide by the total number of symptoms experienced. Based on a reanalysis of the severity data using the desired score calculation, Estrasorb meets the efficacy criteria.

1.2 OVERVIEW OF CLINICAL PROGRAM AND STUDIES REVIEWED

Study E99-1 is a randomized, double-blind, placebo-controlled, multicenter, parallel group study. The objective was to compare Estrasorb to placebo during a 12-week treatment period. Patients were postmenopausal women who experienced a mean of at least 60 moderate to severe hot flushes during a 2-week baseline screening period. A total of 200 women were randomized, using a 1:1 ratio, to the two treatment groups. Randomization was done within stratum, defined as intact uterus or hysterectomy. Efficacy data was collected using a daily diary of the frequency and severity of hot flushes, recorded throughout the 2-week baseline screening and treatment periods.

Study E99-1 is an appropriately designed Phase III study for the indication of treatment of vasomotor symptoms in postmenopausal women. The sample size was sufficient for this to be the single Phase III study to support this NDA.

1.3 PRINCIPAL FINDINGS

Estrasorb was statistically significantly better than placebo at both Week 4 and Week 12 for the symptom severity score (all p-values ≤ 0.001).

2 STATISTICAL REVIEW AND EVALUATION OF EVIDENCE

2.1 INTRODUCTION AND BACKGROUND

The indication of treatment of vasomotor symptoms in postmenopausal women requires one placebo-controlled Phase III confirmatory trial of 12 weeks treatment duration. Subjects are postmenopausal women with at least 56 moderate-to-severe vasomotor symptoms (MSVS) per week during a 2-week baseline screening period. There are four co-primary endpoints, and a new product should be significantly better than placebo for all four. These endpoints are:

- Mean change in number of MSVS from baseline to Week 4
- Mean change in number of MSVS from baseline to Week 12
- Mean change in severity of vasomotor symptoms from baseline to Week 4
- Mean change in severity of vasomotor symptoms from baseline to Week 12

In the analysis of the original NDA submission, severity of vasomotor symptoms was assessed using a severity index. In this review, the severity is being reanalyzed using a severity score. The severity index was calculated a s a weighted sum of vasomotor symptoms, with mild receiving a weight of 1, moderate a weight of 2, and severe a weight of 3. The index did not divide by the total number of vasomotor symptoms experienced by a patient. The only difference in the calculation of the severity score in this review is that it does divide by the total number of symptoms experienced. The goal is to standardize the way severity of vasomotor symptoms is evaluated for this indication.

2.2 DATA ANALYZED AND SOURCES

Study E99-1 is a randomized, double-blind, placebo-controlled, multicenter, parallel arm trial. The primary objective of this study was to compare the efficacy of Estrasorb (7.5 mg estradiol) to placebo. It was designed as the single Phase III confirmatory trial for the indication of treatment of vasomotor symptoms in postmenopausal women. The treatment duration was 12 weeks.

Patients were postmenopausal women who had a mean of at least 60 moderate-to-severe hot vasomotor symptoms (MSVS) per week during the two-week baseline screening period. After confirming eligibility during the baseline screening period, the women were randomized at a 1:1 ratio to the two treatment groups. A total of 200 women, 100 per group, were enrolled. Randomization was stratified by hysterectomy status (intact uterus vs. hysterectomy) to ensure balance across the groups. Subjects recorded vasomotor symptoms, both frequency and severity, on daily diary cards during the baseline and treatment periods.

The primary analysis, as planned in the protocol, was ANCOVA model with mean change from baseline in the number of MSVS as the dependent variable, with baseline mean as the covariate, and factors for treatment and strata. The term for the treatment-by-strata

interaction would be tested but would be dropped from the model for the primary analyses if not significant (p-value > 0.05).

The primary efficacy analyses planned in the protocol were the change from baseline in the number of MSVS at Week 4 and Week 12. The protocol planned for the analyses of the severity of all vasomotor symptoms at Week 4 and Week 12 as secondary analyses. All four analyses would use the same ANCOVA model. The Intent-to-Treat (ITT) patient population was appropriately defined as all subjects randomized with at least one on-treatment measurement.

2.3 STATISTICAL EVALUATION OF EVIDENCE ON EFFICACY

The ANCOVA model planned by the sponsor is appropriate for the analyses. This model includes terms for baseline, treatment, and stratum. The treatment-by-stratum interaction term will be tested and dropped from the model if not significant.

Two ANCOVA models are reported, one for each of the severity score timepoints. Since Estrasorb must be statistically significantly better than placebo on both endpoints, no adjustment for multiplicity is needed. The sponsor's analysis used severity index. My reanalysis here uses the severity score calculated as the severity index divided by the number of symptoms recorded.

Table 1 on the following page shows the results for the severity score endpoints. The between-group comparisons are only done at Week 4 and Week 12. Week 8 results are shown for consistency with other products for this indication, but between-group comparisons are not performed at Week 8.

The comparisons for the severity score endpoint at the Week 4 and Week 12 timepoints indicate that Estrasorb is significantly different from, and better than, placebo. These results support the efficacy of Estrasorb for the treatment of vasomotor symptoms in postmenopausal women.

APPEARS THIS WAY
ON ORIGINAL

Table 1: Reviewer's Results – Severity Score of Vasomotor Symptoms (ITT)

Mean Severity Score of Vasomotor Symptoms		Estrasorb (N=100)	Placebo (N=100)
Baseline	N	100	100
	Mean	2.36	2.44
	Std. Dev.	0.36	0.37
Change from Baseline to Week 4	N	96	97
	Mean	-0.89	-0.45
	Std. Dev.	1.04	0.75
	p-value*	0.0010	
Change from Baseline to Week 8	N	91	94
	Mean	-1.32	-0.63
	Std. Dev.	1.12	0.88
Change from Baseline to Week 12	N	90	90
	Mean	-1.44	-0.55
	Std. Dev.	1.04	0.91
	p-value*	<0.0001	

Source: SAS datasets
* P-values for the treatment group comparisons are from an ANCOVA model with a term for treatment with baseline severity as the covariate. The dependent variable is the change from baseline. The between-group comparison is done for the primary endpoints at Week 4 and Week 12, not at Week 8.

2.4 CONCLUSIONS AND RECOMMENDATIONS

The results of study E99-1 support the efficacy of Estrasorb for the treatment of vasomotor symptoms. This design of this study was appropriate to assess efficacy for this indication, and my original review showed that Estrasorb was statistically significantly different from placebo on the change in frequency at week 4 and week 12 endpoints. This review shows that Estrasorb is statistically significantly different from placebo on the severity score at week 4 and week 12. Therefore Estrasorb has met the efficacy criteria for all four primary efficacy endpoints (all p-values <= 0.0010).
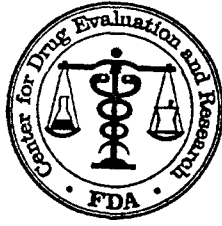
APPEARS THIS WAY
ON ORIGINAL

/s/

Katherine Meaker
7/9/03 03:40:13 PM
BIOMETRICS

Mike Welch
7/10/03 09:17:45 AM
BIOMETRICS
Concur with review

APPEARS THIS WAY
ON ORIGINAL

Statistical Review and Evaluation

CLINICAL STUDIES

NDA:	21-371
Name of drug:	Estrasorb (estradiol ⸺
Applicant:	Novavax, Inc.
Indication:	⸺
Documents reviewed:	Vols. 1.001, 1.025-1.051, 1.053;
	\\CDSESUB1\N21371\N_000\2001-06-29˙
Project manager:	Dornette Spell-LeSane (HFD-580)
Clinical reviewer:	Theresa van der Vlugt, M.D. (HFD-580)
Dates:	Received 6/29/01; User fee (10 month) 4/29/02; HFD-580 requested date 4/8/02
Statistical reviewer:	Kate Meaker, M.S. (HFD-715)
Statistics team leader:	Mike Welch, Ph.D. (HFD-715)
Biometrics division director:	Ed Nevius, Ph.D. (HFD-715)

**APPEARS THIS WAY
ON ORIGINAL**

1 EXECUTIVE SUMMARY OF STATISTICAL FINDINGS

1.1 CONCLUSIONS AND RECOMMENDATIONS

This NDA included a single well-controlled, Phase III, clinical trial to support the efficacy of Estrasorb for the treatment of vasomotor symptoms in postmenopausal women. That study, E99-1, was appropriately designed to assess efficacy for this indication. The results indicate that Estrasorb is statistically significantly different from placebo for all four primary efficacy endpoints. In my opinion there is sufficient evidence to support the efficacy of Estrasorb for this indication.

1.2 OVERVIEW OF CLINICAL PROGRAM AND STUDIES REVIEWED

Study E99-1 is a randomized, double-blind, placebo-controlled, multicenter, parallel group study. The objective was to compare Estrasorb to placebo during a 12-week treatment period. Patients were postmenopausal women who experienced a mean of at least 60 moderate to severe hot flushes during a 2-week baseline screening period. A total of 200 women were randomized, using a 1:1 ratio, to the two treatment groups. Randomization was done within stratum, defined as intact uterus or hysterectomy. Efficacy data was collected using a daily diary of the frequency and severity of hot flushes, recorded throughout the 2-week baseline screening and treatment periods.

Study E99-1 is an appropriately designed Phase III study for the indication of treatment of vasomotor symptoms in postmenopausal women. The sample size was sufficient for this to be the single Phase III study to support this NDA.

1.3 PRINCIPAL FINDINGS

There are 4 co-primary endpoints for this indication. Estrasorb was statistically significantly better than placebo for all four endpoints (all p-values ≤ 0.0005). There was no notable relationship between treatment effect and stratum (intact uterus vs. hysterectomy) or body mass index (BMI).

2 STATISTICAL REVIEW AND EVALUATION OF EVIDENCE

2.1 INTRODUCTION AND BACKGROUND

The indication of treatment of vasomotor symptoms in postmenopausal women requires one placebo-controlled Phase III confirmatory trial of 12 weeks treatment duration.

3

Subjects are postmenopausal women with at least 56 moderate-to-severe vasomotor symptoms (MSVS) per week during a 2-week baseline screening period. There are four co-primary endpoints, and a new product should be significantly better than placebo for all four. These endpoints are:

- Mean change in number of MSVS from baseline to Week 4
- Mean change in number of MSVS from baseline to Week 12
- Mean change in severity of vasomotor symptoms from baseline to Week 4
- Mean change in severity of vasomotor symptoms from baseline to Week 12

2.2 DATA ANALYZED AND SOURCES

Study E99-1 is a randomized, double-blind, placebo-controlled, multicenter, parallel arm trial. The primary objective of this study was to compare the efficacy of Estrasorb (7.5 mg estradiol) to placebo. It was designed as the single Phase III confirmatory trial for the indication of treatment of vasomotor symptoms in postmenopausal women. The treatment duration was 12 weeks.

Patients were postmenopausal women who had a mean of at least 60 moderate-to-severe hot vasomotor symptoms (MSVS) per week during the two-week baseline screening period. After confirming eligibility during the baseline screening period, the women were randomized at a 1:1 ratio to the two treatment groups. A total of 200 women, 100 per group, were enrolled. Randomization was stratified by hysterectomy status (intact uterus vs. hysterectomy) to ensure balance across the groups. Subjects recorded vasomotor symptoms, both frequency and severity, on daily diary cards during the baseline and treatment periods.

The primary analysis, as planned in the protocol, was ANCOVA model with mean change from baseline in the number of MSVS as the dependent variable, with baseline mean as the covariate, and factors for treatment and strata. The term for the treatment-by-strata interaction would be tested but would be dropped from the model for the primary analyses if not significant (p-value > 0.05).

The primary efficacy analyses planned in the protocol were the change from baseline in the number of MSVS at Week 4 and Week 12. The protocol planned for the analyses of the mean severity of all vasomotor symptoms at Week 4 and Week 12 as secondary analyses. All four analyses would use the same ANCOVA model. The Intent-to-Treat (ITT) patient population was appropriately defined as all subjects randomized with at least one on-treatment measurement.

2.3 STATISTICAL EVALUATION OF EVIDENCE ON EFFICACY

The ANCOVA model planned by the sponsor is appropriate for the analyses. This model includes terms for baseline, treatment, and stratum. The treatment-by-stratum interaction term will be tested and dropped from the model if not significant.

The sponsor planned to analyze the mean change in number of MSVS at Week 4 and Week 12 as the primary efficacy endpoints. The analyses of the mean change in severity at Week 4 and Week 12 were planned in the protocol as secondary endpoints. However, all four are desired as co-primary endpoints by DRUDP and will be considered as equally important in this review.

A total of four ANCOVA models are reported, one for each of the co-primary endpoints. Since Estrasorb must be statistically significantly better than placebo on all four endpoints, no adjustment for multiplicity is needed. My analyses, results, and conclusions match those of the sponsor, so they will only be presented once.

A comparison of the number and reasons for dropouts in the study indicated there was no notable differences between the groups in the disposition of subjects. Overall only 9% of the subjects did not complete the 12-week treatment period.

Tables 1 and 2 show the results for the four co-primary endpoints. The between-group comparisons are only done at Week 4 and Week 12. Week 8 results are shown for consistency with other products for this indication, but between-group comparisons are not performed at Week 8.

Table 1: Reviewer's Results – Frequency of Moderate-to-Severe Vasomotor Symptoms (ITT)

Mean Number of MSVS		Estrasorb (N=100)	Placebo (N=100)
Baseline	N	100	100
	Mean	13.05	13.63
	Std. Dev.	5.78	5.48
Change from Baseline to Week 4	N	96	97
	Mean	-8.56	-5.97
	Std. Dev.	6.19	4.76
	p-value*	0.0002	
Change from Baseline to Week 8	N	91	94
	Mean	-10.74	-6.67
	Std. Dev.	6.99	5.26
Change from Baseline to Week 12	N	90	90
	Mean	-11.11	-7.20
	Std. Dev.	6.84	5.39
	p-value*	<0.0001	

Source: SAS datasets
* P-values for the treatment group comparisons are from an ANCOVA model with terms for treatment and stratum (hysterectomy vs. intact uterus) with baseline frequency as the covariate. The between-group comparison is done for the primary endpoints at Week 4 and Week 12, not at Week 8.

Table 2: Reviewer's Results – Severity of Vasomotor Symptoms (ITT)

Mean Severity of Vasomotor Symptoms		Estrasorb (N=100)	Placebo (N=100)
Baseline	N	100	100
	Mean	33.17	35.55
	Std. Dev.	16.93	16.11
Change from Baseline to Week 4	N	96	97
	Mean	-21.32	-15.62
	Std. Dev.	16.55	12.30
	p-value*	0.0005	
Change from Baseline to Week 8	N	91	94
	Mean	-27.11	-17.18
	Std. Dev.	19.59	13.83
Change from Baseline to Week 12	N	90	90
	Mean	-27.95	-18.44
	Std. Dev.	19.38	14.52
	p-value*	<0.0001	

Source: SAS datasets

* P-values for the treatment group comparisons are from an ANCOVA model with terms for treatment and stratum (hysterectomy vs. intact uterus) with baseline severity as the covariate. The between-group comparison is done for the primary endpoints at Week 4 and Week 12, not at Week 8.

The comparisons for all four of the co-primary endpoints indicate that Estrasorb is significantly different from, and better than, placebo. These results support the efficacy of Estrasorb for the treatment of vasomotor symptoms in postmenopausal women.

2.4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

Strata

Randomization was done within stratum (intact uterus vs. hysterectomy) to ensure balance across the treatment groups. In the protocol the sponsor planned an analysis to test for an interaction between treatment and stratum using an ANCOVA model. If the interaction term was not significant at the 0.05 level, the interaction term would be dropped from the model.

In the analyses of the four primary endpoints, the results for the treatment-by-stratum interaction term indicated a possible relationship (p-value = 0.0730) for the mean change in number of MSVS at Week 4, but not for the other three endpoints. Further investigation of the results showed that the mean change was larger for both strata in the Estrasorb group

than in the placebo group. The magnitude of the change was slightly larger in the intact uterus stratum compared to the hysterectomy stratum. The interaction was not driving the overall treatment results and dropping the interaction term from the model did not change the results for any of the endpoints. Therefore the interaction term was dropped from the primary analyses.

Site

There were no apparent differences in the results across the 18 sites that enrolled patients. The two largest sites (#3 had 23% of total patients; #1 had 15%) had no notable differences from the overall results, and no site was driving the results.

Body Mass Index (BMI)

The Medical Officer requested an analysis by body mass index (BMI) because of a possible imbalance across the two treatment groups as baseline. In the placebo treatment group, 36% of the subjects had BMI >27 kg/m², while in the Estrasorb group 50% of the subjects had BMI >27 kg/m². This is an exploratory analysis only.

I used an ANCOVA model with terms for baseline, treatment, BMI category (≤ 27, >27), and treatment-by-BMI interaction for each of the four primary endpoints. None of the analyses indicated any relationship between BMI and treatment effect.

2.5 CONCLUSIONS AND RECOMMENDATIONS

The results of study E99-1 support the efficacy of Estrasorb. This design of this study was appropriate to assess efficacy for this indication, and Estrasorb was statistically significantly different from placebo for all four primary efficacy endpoints (all p-values ≤ 0.0005).

2.6 LABELING COMMENTS

The sponsor presents the result for the change from screening for the frequency of MSVS is Table 7.0 of the label. As proposed, the table includes Screening (2-week baseline screening), Week −1 (placebo run-in), Week 1, Week 4, Week 8, and Week 12 time points. I suggest that [   ] to be consistent with other labels for this indication. Also, the P-value column should only list values for Week 4 and Week 12 since these are the time points at which hypothesis tests are done. The sponsor notes ND = Not Done for Screening, and I suggest the same notation be used for Week 8.

/s/

Katherine Meaker
4/22/02 02:27:29 PM
BIOMETRICS

Mike Welch
4/22/02 04:00:51 PM
BIOMETRICS
Concur with review

S. Edward Nevius
4/22/02 04:17:59 PM
BIOMETRICS
Concur with review.

**APPEARS THIS WAY
ON ORIGINAL**

DRAFT GUIDANCE FOR LABELING FOR COMBINED ORAL
CONTRACEPTIVES DATED OCTOBER 2003 CAN BE FOUND
ON THE FDA WEBSITE
(http://www.fda.gov/cder/guidance/index.htm)

27 pages redacted