CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER: NDA 19726/S18

STATISTICAL REVIEW(S)

Statistical Review and Evaluation
Clinical Studies

Date: JUN 15 1997

NDA #: 19-726 / S018

Applicant: ZENECA

Name of Drug: Zoladex (goserelin acetate implant)

Indication: endometrial thinning prior to endometrial ablation

Documents Reviewed: Vol. 18.1; 18.22-18.30

Statistical Reviewer: Kate Meaker, M.S. (HFD-715)

Medical Input: Heidi Jolson, M.D. (HFD-580)

Summary of Studies

ZOLADEX is an agonist intended for gonadal suppression. The desired indication for this submission is for use as an endometrial thinning agent prior to endometrial ablation.

Study 9393IL/0022 (0022) was a randomized, double-blind, placebo-controlled, multicenter study. The sample consisted of pre-menopausal women, age 30 or older with regular menstrual cycles, who had been diagnosed with dysfunctional uterine bleeding (DUB) and selected for total endometrial ablation. After screening to determine eligibility, patients were randomly assigned by a 1:1 ratio to either the Zoladex 3.6 mg or the Sham (placebo) treatment group. The treatment regimen consisted of receiving 2 treatment injections, 28 days apart, with surgery for endometrial ablation performed 6 weeks +/- 3 days after the first injection. The timing of the first injection was based on the length of the patient's menstrual cycle, with the goal that the endometrium would be at its thinnest stage (day 7 of menstrual cycle) at the time of surgery for all subjects. Surgery was performed primarily by loop resection. However, rollerball could be used in the cornual regions of the uterus where the myometrium is very thin (Vol. 18.23, pg. 27). Efficacy variables were measured at the time of surgery and at the follow-up visit 24 weeks after surgery.

Study 9393IP/0003 (0003) was an open-label, randomized, single-center controlled clinical trial. The study population consisted of pre-menopausal women, age 20 or older, who had been diagnosed with dysfunctional uterine bleeding (DUB) and selected for total endometrial ablation. After screening to determine eligibility, patients were randomly assigned on an equal basis to 4 treatment groups. Two active-control treatment arms, each containing Danazol, were included in the clinical trial but are not being considered as a comparator in this submission (agreement in pre-NDA mtg. Oct. 26, 1995)[1]. The only treatment arms of interest in this submission are the 2 groups which received Zoladex. One group received 1 injection of Zoladex (3.6 mg) with endometrial laser ablation (ELA) surgery scheduled for 4 weeks after the injection. The other treatment group received 2 injections of Zoladex (3.6 mg), 28 days apart, with ELA surgery scheduled for 4 weeks after the second injection. For both treatment groups, the first injection was scheduled to occur during menstruation, and a follow-up visit was scheduled for 24 weeks after surgery.

[1] Danazol is considered investigational for this indication in the United States. Therefore it was not considered to be a suitable positive control.

The main differences between these 2 studies are:

- In study #0022 loop resection was the primary surgical procedure, while in study #0003 the procedure was laser ablation.
- The timing of the first injection: (day 7 of menstrual cycle in study #0022; during menstruation in study #0003)
- The timing of surgery: (2 weeks after 2nd injection in study #0022; 4 weeks after last injection in study #0003)
- Study #0022 included a placebo (sham) control group; study #0003 used no comparator for the analysis.

Table 1: Summary of Randomized, Controlled Clinical Studies

Study Number (Dates Conducted)	Number of Centers (Locations)	Total Sample Size	Type of Control	Design	Duration of Treatment
9393IL/0022 (6/94 - 11/95)	37 (all non-U.S.)	ZOLADEX (n=180) Sham (n=178)	Placebo	randomized, double-blind, multicenter, 2 parallel trmt groups	8 weeks; 2 injections, 28 days apart, surgery 6 wks after 1st injection; Follow-up 12 and 24 wks post-surgery
9393IP/0003 (3/92 - 5/94)	1 (U.K.)	n=40 each group:* ZOLADEX, 4 wks. ZOLADEX, 8 wks. Danazol, 4 wks. Danazol, 8 wks.	Active (but Danazol trmt gps not included in analysis.)	open-label, randomized, 4 parallel trmt groups	4 or 8 weeks; surgery 28 days after last trmt injection; Follow-up 24 wks post-surgery

* For study 9393IP/0003, only the two ZOLADEX treatment groups are included in analyses (n=80).

STUDY # 9393IL/0022

Background

Study 9393IL/0022 (0022) was a randomized, double-blind, placebo-controlled, multicenter study. The study population consisted of pre-menopausal women, age 30 or older with regular menstrual cycles, who had been diagnosed with dysfunctional uterine bleeding (DUB) and selected for total endometrial ablation. Subjects must have had no desire to have further children. After screening to determine eligibility, patients were randomly assigned by a 1:1 ratio to either the Zoladex 3.6 mg or the Sham (placebo) treatment group.

The treatment regimen consisted of receiving 2 treatment injections, 28 days apart, with surgery for endometrial ablation performed 6 weeks +/- 3 days after the first injection. The timing of the first injection was based on the length of the patient's menstrual cycle, with the goal that the endometrium would be at its thinnest stage (day 7 of menstrual cycle) at the time of surgery for all subjects. Surgery was performed primarily by loop resection. However, rollerball could be used in the cornual regions of the uterus where the myometrium is very thin (Vol. 18.23, pg. 27).

The study intended to include 45 centers, but only 37 centers recruited patients to participate.

Primary variables of interest specified by the applicant were:

Incidence of amenorrhea (24 weeks post-surgery)
Endometrial thickness (immediately before surgery)
Change in Blood Loss Score (24 weeks post-surgery vs. pre-surgery)

In discussions with the medical officers, it was decided that only incidence of amenorrhea and endometrial thickness will be used for the regulatory decision.

Secondary variables of interest were:

Duration of surgery
Ease of surgery

A total of 358 patients were randomized to the 2 treatment groups. The 2 groups were similar with regard to most of the demographic characteristics at baseline, as shown in Table 2. However, the treatment groups were significantly different for the baseline weight (p=.0004).

Table 2: Demographic characteristics (Study #0022)

	Zoladex (n=180)		Sham (n=178)	
Characteristic	Mean	S.D.	Mean	S.D.
Age (years)	41.4	4.7	41.2	5.2
Weight (kg)	71.8	15.6	66.3	13.5
Uterine Cavity Length (cm)	6.9	1.7	6.5	1.6
Average Duration of Menstrual Cycle (days)	27.0	3.4	27.5	3.3
	n	%	n	%
Race				
Caucasian	173	96.1	173	97.2
Other	7	3.9	5	2.8

Source: Vol. 18.23, Table 5

APPEARS THIS WAY
ON ORIGINAL

The disposition of the subjects in the 2 treatment groups was similar in terms of both the number of drop-outs at any stage and the reason for discontinuation (see Tables 3 & 4).

Table 3: Disposition of subjects by group (Study #0022)

	Zoladex		Sham	
	n	rand. %	n	rand. %
Randomized	180	100.0	178	100.0
Received treatment	180	100.0	177	99.4
Received surgery	177	98.3	175	98.3
Completed trial	174	96.7	172	96.6

Source: Vol. 18.23, Table 7

Table 4: Reasons for Discontinuation (Study #0022)

	Zoladex		Sham	
	n	% rand.	n	% rand.
Withdrew before treatment				
Protocol non-compliance	0	0.0	1	0.6
Withdrew during trmt; before surgery				
Adverse event	2	1.1	1	0.6
Other	1	0.6	1	0.6
Withdrew at or after surgery				
Adverse event	1	0.6	2	1.1
Protocol non-compliance	0	0.0	1	0.6
Other	2	1.1	0	0.0

Source: Vol. 18.23, Table 7

Applicant's Analysis

The efficacy analyses reported by the applicant were conducted on a modified intent-to-treat (ITT) basis. The true ITT (all randomized subjects) was 'modified' in the sense that patients with no data for an endpoint were excluded in the analysis of that endpoint. All randomized subjects who had data for an endpoint were included in the analysis of that endpoint. (Vol. 18.23, pg. 41)

The efficacy variables were measured on different scales which required different methods of analysis. For each variable, the measurement scale, applicant's analysis method, and applicant's results are described below.

Incidence of Amenorrhea (primary efficacy variable)

The incidence of amenorrhea at 24 weeks post-surgery is a binary variable. The logistic regression model, with factors for treatment and center, was used by the applicant. The results are reported as the odds ratio with a 95% confidence interval (Table 5a). An odds ratio greater than 1 favors Zoladex over the Sham treatment. The conclusion is that the estimated odds of experiencing amenorrhea 24 weeks post-surgery in the Zoladex group were twice (2.11) those in the Sham group and that this is a statistically significant difference. The lower bound of the 95% confidence interval for the odds ratio is greater than 1, also favoring Zoladex.

Table 5a: Applicant's Results (Modified ITT) (Study #0022)

	Zoladex (n=175)		Sham (n=171)		Logistic Regression Results	
	n	%	n	%	Odds Ratio	95% Confidence Interval
Incidence of Amenorrhea (%)	70	40.0	44	25.7	2.11 (p-value=.0039)	(1.27, 3.50)

Source: Vol. 18.23, Table 11.

Endometrial Thickness (primary efficacy variable)

Endometrial thickness immediately pre-surgery was measured on a continuous scale (mm). In the study report, the applicant notes that the data were skewed, and therefore a geometric mean transformation method was applied for the analysis. The geometric mean is the nth root of the product of the data, which reduces the weight given to outliers in the original data. This approach transforms the variable from the original scale to the log scale. The mean of the log-transformed data equals the log of the geometric mean. The treatment group comparison is a ratio of the geometric means instead of a difference. After transforming the endometrial thickness response variable to the log scale, an ANOVA model with terms for treatment and center was used. The model did not include the treatment-by-center interaction term, which was not specified in the protocol. The results of the applicant's analysis using the geometric mean model are shown in Table 5b.

A value for the ratio less than 1 favors Zoladex. The applicant concluded the endometrium was statistically significantly thinner in the Zoladex group than in the Sham group (p-value=.0001). The estimated ratio shows that the average thickness in the Zoladex group was 46% of that in the Sham group.

Table 5b: Applicant's Results (Modified ITT) (Study #0022)

	Zoladex		Sham		Geometric Means Model Results	
	n	Adjusted * geometric mean (Std. Err)	n	Adjusted * geometric mean (Std. Err)	Ratio	95% Confidence Interval on Ratio
Endometrial Thickness (mm)	176	1.61 (1.05)	173	3.53 (1.05)	0.46 (p-value=.0001)	(0.41, 0.52)

Source: Vol. 18.23, Table 19.
* Adjusted for center in ANOVA model.

Change in Total Blood Loss (primary efficacy variable)

This efficacy variable is also referred to as change in menstrual pattern. This is calculated using pre and post-operation menstrual scores derived from menstrual diaries. The pre-surgery diary was completed during the screening phase (before randomization) and the post-surgery diary was completed at 24 weeks post-surgery. Subjects with amenorrhea at 24 weeks post-surgery were assigned a value of zero for the post-surgery score. The change is calculated as the difference (post minus pre) of the 2 scores for each subject. Therefore a negative value indicates a positive result, reduction in total blood loss.

The analysis described in the protocol was an Analysis of Covariance (ANCOVA) model with factors for treatment and center along with the pre-treatment score as a covariate. The data for the change in total blood loss were skewed, so the applicant used a non-parametric method. For the study report, the change in total blood loss values were transformed to ranks, and the same ANCOVA model proposed in the protocol was applied to the ranked data instead of the original data. The results appear in Table 5c. The applicant's conclusion is that there is not sufficient evidence of a difference between the treatment groups with respect to the median change in total blood loss score.

Table 5c: Applicant's Results (Modified ITT) (Study #0022)

	Zoladex		Sham		ANCOVA on Rank Transformed Data Results	
	n	Adjusted * median	n	Adjusted * median	Difference in Adjusted Medians	95% Confidence Interval
Change in Total Blood Loss Score	174	-306.0	169	-305.5	-0.5 (p-value=.91)	(-8.49, 7.57)

Source: Vol. 18.23, Table 16.
* Adjusted for center and baseline blood loss score in the ANCOVA model.

Duration of Surgery (secondary efficacy variable)

Duration of surgery was measured on a continuous scales (minutes), and the data were skewed toward the higher values. In the study report, the applicant used the same geometric means (log transform) method which was applied to the endometrial thickness variable. The observed values for duration were transformed to the log scale, and then an ANCOVA model, which included terms for treatment, center, and baseline uterine length was used for the analysis. The results, using the ratio for comparison, are shown in Table 5d.

The estimate of the ratio is less than 1, as is the upper bound of the 95% confidence interval for the ratio. A ratio less than 1 favors Zoladex as having a shorter duration of surgery than the Sham treatment. The applicant concluded that the duration of surgery was significantly shorter in the Zoladex group than in the Sham group.

Table 5d: Applicant's Results (Modified ITT) (Study #0022)

	Zoladex		Sham		Geometric Means Model Results	
	n	Adjusted * geometric mean (Std. Err)	n	Adjusted * geometric mean (Std. Err)	Ratio	95% Confidence Interval on Ratio
Duration of Surgery (min)	174	18.22 (1.03)	174	23.30 (1.03)	0.78 (p-value=.0001)	(0.73, 0.84)

Source: Vol. 18.23, Table 23.
* Adjusted for center and baseline uterine length in the ANCOVA model.

Ease of Surgery (secondary efficacy variable)

Ease of surgery was measured on a 3-point ordinal scale. The categories were: easier than usual, routine procedure, and more difficult than usual. This was assessed by the surgeon immediately after the surgery was completed. No objective criteria for this assessment were listed in the protocol or study report.

This was analyzed using the continuation odds model, which is appropriate for ordinal data. This model provides an odds ratio for one category of interest compared to the other categories combined. In this case, the odds ratio is the odds of having easier than usual surgery compared to routine or more difficult surgery for the Zoladex treatment group versus the Sham treatment group. The results (Table 5e) indicate a statistically significant treatment difference (p-value=.0001) which favors Zoladex. The estimated odds ratio, adjusted for center in the model, indicates that the odds of having easier surgery are more than 6 times greater after receiving Zoladex than after receiving the Sham treatment.

Table 5e: Applicant's Results (Modified ITT) (Study #0022)

	Zoladex (n=177)		Sham (n=175)		Continuation Odds Model Results	
	n	%	n	%	Odds Ratio*	95% Confidence Interval
Ease of Surgery					6.45 (p-value=.0001)	(4.14, 10.04)
Easier than usual	74	41.8	15	8.6		
Routine	82	46.3	89	50.9		
More difficult	21	11.9	71	40.6		

Source: Vol. 18.23, Table 24.
* Odds ratio estimate adjusted for center in logistic regression model.

Reviewer's Analysis

The modified intent-to-treat population of subjects which the applicant used includes all randomized subjects who had data for each variable. This same population is used for this reviewer's analyses. There were few missing observations, with at most 2 subjects missing in either treatment group in any of the analyses.

Incidence of Amenorrhea (primary efficacy variable)

In the protocol for study #0022 (Vol. 18.24, pg. 108) the sample size was determined on differences in the amenorrhea rates. In the NDA, however, the applicant calculated amenorrhea rates, but based the conclusion using the odds ratio results. This reviewer calculated the difference between the amenorrhea rates to confirm the applicant's results using the method specified in the protocol (Table 6a). The difference in rates (14.3) favors Zoladex. The lower bound of the 95% confidence interval is greater than zero, which indicates the difference in amenorrhea rates is statistically significant.

In the protocol, the clinically meaningful difference was defined as a 20% difference (Vol. 18.24, pg. 108). The confidence interval on the difference in rates does not exclude this meaningful difference. Thus the results indicate a significant difference, but it may not be a clinically meaningful difference.

Table 6a: Reviewer's Results (Modified ITT) (Study #0022)

	Zoladex (n=175)		Sham (n=171)		Difference in Rates	
	n	%	n	%	Difference	95% Confidence Interval
Incidence of Amenorrhea (%)	70	40.0	44	25.7	14.3	(4.5, 24.0)

Endometrial Thickness (primary efficacy variable)

The analysis for endometrial thickness proposed in the protocol (Vol. 18.24, pg. 110) was the Mann-Whitney U-test estimating the difference between the treatment group medians. This is a non-parametric test and is appropriate for skewed data. A difference of 1.2 mm was specified in the protocol as the clinically meaningful difference.

The applicant did not use the proposed non-parametric method in the study analysis, but instead compared the geometric means. The treatment group comparison from this approach was in the form of a ratio of the geometric means, which is not on a comparable scale as the defined clinically meaningful difference. This reviewer used a non-parametric test, the Wilcoxon Rank-Sum test (equivalent to the Mann-Whitney U-test), on the untransformed data to confirm the applicant's conclusions and to provide a confidence interval on the medians from the untransformed data. This model does not adjust for center. The results are presented in Table 6b.

Note that the arithmetic difference between the medians for the 2 groups does not necessarily equal the median of the differences estimated by the Wilcoxon method. This is because the Wilcoxon method calculates the difference between all possible pairs of subjects ($[n_1 \times n_2] \div 2$ = total # of pairs) and then calculates the median of all differences, not the difference of the 2 group medians.

Table 6b: Reviewer's Results (Modified ITT) (Study #0022)

	Zoladex		Sham		Wilcoxon Rank-Sum Results	
	n	Median (25th, 75th p-tile)	n	Median (25th, 75th p-tile)	Median Difference	95% Confidence Interval
Endometrial Thickness (mm)	176	1.50 (1.00, 2.50)	173	3.55 (2.50, 5.00)	-1.85 (p-value=.0001)	(-2.2, -1.5)

The results indicate that there is a statistically significant difference in the median between the treatment groups. The lower bound of the 95% confidence interval is greater than 1.2 mm, which exceeds the clinically meaningful difference defined in the protocol and favors Zoladex.

Change in Total Blood Loss (primary efficacy variable)

The applicant specified this as a primary efficacy variable for this study in the protocol (Vol. 18.24, pg 110). In discussions with the medical officers, it was decided that the change in blood loss variable was not of primary concern for the regulatory decision.

The medical officers felt that the applicant's results for the change in total blood loss score (no treatment difference) contradicts the results for incidence of amenorrhea (significant treatment difference). The applicant had used an ANCOVA model, but had transformed both the response variable, change (post-pre), and the covariate (pre-surgery blood loss score) to ranks prior to the analysis. The ANCOVA model is:

Change (Post - Pre) = Treatment + Center + Pre-surgery blood loss score

The analysis discussed in the protocol was this same ANCOVA model, but using the data on the original scale rather than transforming the variables to ranks.

First this reviewer considered the distributions for the change in blood loss score, pre-surgery score, and post-surgery score to investigate why the rank transformation was used by the applicant. Subjects who had amenorrhea at 24 weeks post-surgery did not complete a menstrual diary and were assigned a value of zero for the post-surgery blood loss score. This is a logical value to assign, but it impacts the results of the ANCOVA model proposed by the applicant to analyze change in blood loss. For subjects with amenorrhea, there is a perfect linear relationship [-Y = pre-surgery score] between the change and the covariate variable in the original data. All the variation in Y (change) would be explained by the covariate for the amenorrheic subjects. Only the subjects who were not amenorrheic would contribute to the unexplained variation in Y in this model, but the degrees of freedom would reflect the total number of subjects. In the rank scale this relationship is not true. However the reason the applicant gave for using the rank transformation was that the data was skewed, not because of the impact of the coding for amenorrheic subjects.

This reviewer considered 2 options for the analysis, both using the data in the original scale rather than the rank transformation used by the applicant. For the intent-to-treat population, the change in blood loss score (post-pre) was modeled with an ANOVA model which included the treatment and center factors but not the pre-surgery blood loss covariate. For the non-amenorrheic subgroup, the original ANCOVA model proposed by the applicant was used. The results for these analyses are shown in Tables 6c and 6d, respectively (next page). The conclusion in both analyses is that there is not a significant treatment difference for change in blood loss score. This reviewer checked histograms of the post-surgery and change (post-pre) in blood loss scores to confirm the results, and found no patterns (i.e. no indication of a treatment effect) in the graphs.

Table 6c: Reviewer's Results (Modified ITT) (Study #0022)

	Zoladex		Sham		ANCOVA Results	
	n	Adjusted * mean	n	Adjusted * mean	Difference (Zol-Sham)	95% Confidence Interval
Change in Total Blood Loss Score	174	-360.6	169	-346.0	-14.6 (p-value=.68)	(-96.6, 67.4)

* Adjusted for center in the ANOVA model.

Table 6d: Reviewer's Results (Non-amenorrheic subjects who completed diary 24 weeks post-surgery) (Study #0022)

	Zoladex		Sham		ANCOVA Results	
	n	Adjusted * mean	n	Adjusted * mean	Difference (Zol-Sham)	95% Confidence Interval
Change in Total Blood Loss Score	104	-334.2	125	-347.3	13.1 (p-value=.35)	(-18.5, 44.7)

* Adjusted for center and baseline blood loss score in the ANCOVA model.

Duration of Surgery (secondary efficacy variable)

The analysis for duration of surgery discussed in the protocol (Vol. 18.24, pg. 110) was an ANCOVA model using the original scale (minutes) with terms for treatment, center, and baseline uterine length The applicant's reported geometric means analysis used the same model terms, but transformed duration to the log scale because the data were skewed. To make the treatment group comparison on the original scale (difference instead of ratio) a non-parametric test is appropriate. This reviewer used the Wilcoxon Rank-Sum test for the comparison in terms of the median difference (Table 6e).

Note that the arithmetic difference between the medians for the 2 groups does not necessarily equal the median of the differences estimated by the Wilcoxon method. This is because the Wilcoxon method calculates the difference between all possible pairs of subjects ($[n_1 \times n_2] \div 2$ = total # of pairs) and then calculates the median of all differences, not the difference of the 2 group medians.

The test concludes that there is a statistically significant treatment difference in the duration of surgery (p-value=.0001). The median difference is greater than zero, and the 95% confidence interval excludes zero. With 95% confidence, the population median duration of surgery is at least 2 and up to 6 minutes shorter for the Zoladex treatment group than the Sham treatment group. These results favor Zoladex-treated subjects as having a shorter duration of surgery than the subjects who received the Sham treatment

Table 6e: Reviewer's Results (Modified ITT) (Study #0022)

	Zoladex		Sham		Wilcoxon Rank-Sum Results	
	n	Median (25th, 75th p-tile)	n	Median (25th, 75th p-tile)	Median Difference	95% Confidence Interval
Duration of Surgery (min)	175	17 (12, 25)	175	22 (15, 30)	-4 (p-value=.0001)	(-6, -2)

Ease of Surgery (secondary efficacy variable)

The protocol for study #0022 did not list ease of surgery as an efficacy variable, nor was the analysis method discussed. The continuation odds model used by the applicant is appropriate, so no further analysis was done by this reviewer.

Subgroup Analyses

For descriptive purposes only, this reviewer compared the primary efficacy results for each treatment group by center and age subgroups. There were no trends or patterns among the centers.

For the age comparison, two age groups were defined by this reviewer. The mean age in both treatment groups was between 41 and 42 years, so the subgroups were split at 41 and under; and 42 and older.

For incidence of amenorrhea, the trend within each age group was consistent with the treatment difference found in the main analysis (Table 7a). Specifically, subjects in the Zoladex treatment group had a higher incidence of amenorrhea in both age groups than the subjects in the Sham treatment group. An additional comparison was done by calculating a confidence interval on the differences between the age groups of the treatment group differences within each age group to check for a consistent treatment effect across the age groups. These results are shown below Table 7a, and indicate there is no apparent difference in treatment effect between the age groups because the 95% confidence interval includes zero.

Table 7a: Age Group Comparison (Modified ITT) (Study #0022)

	Age 41 and Under				Age 42 and Older			
	Zoladex (n=83)		Sham (n=84)		Zoladex (n=92)		Sham (n=87)	
	n	%	n	%	n	%	n	%
Incidence of Amenorrhea (%)	30	36.1	23	27.4	40	43.5	21	24.1

Within Age group differences: 8.7% 19.4%
Between Age group difference: -10.7%
Confidence interval on Between Age group diff.: (-23.7% , 45.3%)

For endometrial thickness immediately pre-surgery, the trend across age group was consistent with the treatment difference found in the main analysis (Table 7b). Specifically, subjects in the Zoladex treatment group had a lower endometrial thickness in both age groups than the subjects in the Sham treatment group.

Table 7b: Age Group Comparison (Modified ITT) (Study #0022)

	Age 41 and Under		Age 42 and Older	
	Zoladex (n=85)	Sham (n=84)	Zoladex (n=91)	Sham (n=89)
	Unadjusted Median	Unadjusted Median	Unadjusted Median	Unadjusted Median
Endometrial Thickness (mm)	1.85	3.50	1.50	3.85

The main analysis of the change in blood loss variable did not indicate significant treatment differences using an ANCOVA model which adjusted for baseline value and center. The median values for the subgroup descriptive analysis, shown in Table 7c, do not adjust for baseline value or center. A comparison of treatment effects within each age group shows that subjects who received Zoladex had a greater reduction in blood loss than subjects who received the Sham treatment in both groups. This difference was greater in the 42 and older age group than in the 41 and under age group. These comparisons are descriptive only, ignore the baseline value, and do not imply statistical significance.

Table 7c: Age Group Comparison (Modified ITT) (Study #0022)

	Age 41 and Under		Age 42 and Older	
	Zoladex (n=82)	Sham (n=83)	Zoladex (n=92)	Sham (n=86)
	Median (25th, 75th p-tile)	Median (25th, 75th p-tile)	Median (25th, 75th p-tile)	Median (25th, 75th p-tile)
Change in Total Blood Loss Score	-303.0 (-514.0, -186.0)	-271.0 (-459.0, -182.0)	-367.5 (-550.5, -201.0)	-302.0 (-419.0, -200.0)

Within Age group differences: -32.0 -65.5
Between Age group difference: 33.5

APPEARS THIS WAY
ON ORIGINAL

APPEARS THIS WAY
ON ORIGINAL

Conclusions - Study #0022

This study compares Zoladex to a Sham (placebo) treatment. Of the 3 efficacy variables designated as primary in the protocol, the results for 2 (incidence of amenorrhea and endometrial thickness immediately pre-surgery) indicate a significant treatment effect in the direction which favors Zoladex over the Sham treatment. The 95% confidence interval for the treatment difference (Zoladex - Sham) for endometrial thickness (-2.2 , -1.5) excluded the value of -1.2 mm which was defined as the clinically meaningful difference in the protocol. For incidence of amenorrhea, the significant difference favored Zoladex over the sham treatment, but the 95% confidence interval (4.5, 24.0) did not exclude the clinically meaningful difference of 20% specified in the protocol. Thus the results for incidence of amenorrhea indicate a statistically significant difference, but it may not be a clinically meaningful difference. For the 3rd primary efficacy variable, change in blood loss, there was insufficient evidence to conclude a treatment effect. Two secondary variables related to the duration and ease of the surgical procedure. Significant treatment effects which favored Zoladex were found for both of these variables.

The description of the clinical studies in the proposed label mentions both incidence of amenorrhea and endometrial thickness as primary variables. The results are presented only for incidence of amenorrhea, not for endometrial thickness. The only suggested change to the label is the addition of a heading "Clinical Studies - Endometrial Thinning" to match the format for clinical studies for other indications.

STUDY # 9393IP/0003

Background

Study 9393IP/0003 (0003) was an open-label, randomized, single-center controlled clinical trial, with 4 treatment arms. Two active-control treatment arms, each containing Danazol, were included in the clinical trial but are not being considered as a comparator in this submission (agreement in pre-NDA mtg. Oct. 26, 1995)[2]. The only treatment arms of interest in this submission are the 2 groups which received Zoladex. One group received 1 injection of Zoladex (3.6 mg) with endometrial laser ablation (ELA) surgery scheduled for 4 weeks after the injection. The other treatment group received 2 injections of Zoladex (3.6 mg), 28 days apart, with ELA surgery scheduled for 4 weeks after the second injection. For both treatment groups, the first injection was scheduled to occur during menstruation, and a follow-up visit was scheduled for 24 weeks after surgery.

The study population consisted of pre-menopausal women, age 20 or older, who had been diagnosed with dysfunctional uterine bleeding (DUB) and selected for total endometrial ablation. Subjects could not be pregnant or breast-feeding, and must have had no desire to have further children. After screening to determine eligibility, patients were randomly assigned on an equal basis to the 2 treatment groups. Forty (40) subjects were randomized to each of the treatment groups.

Primary variables of interest were:

Incidence of amenorrhea (24 weeks post-surgery)
Endometrial thickness (immediately before surgery)

Secondary variables of interest were:

Duration of surgery
Difficulty of surgery

[2] Danazol is considered investigational for this indication in the United States. Therefore it was not considered to be a suitable positive control.

A total of 80 patients were randomized to the 2 treatment groups. The 2 groups were similar with regard to the demographic characteristics at baseline, as shown in Table 8.

Table 8: Demographic characteristics (Study #0003)

	Zoladex: 4-week (n=40)		Zoladex: 8-week (n=40)	
Characteristic	Mean	S.D.	Mean	S.D.
Age (years)	42.1	6.2	41.2	6.3
Weight (kg)	67.9	13.8	67.5	15.2
Uterine Cavity Length (cm)	3.55	0.37	3.22	0.31
Average Duration of Menstrual Cycle (days)	7.3	2.9	7.5	2.4

Source: Vol. 18.28, Table T1

The disposition of the subjects in the 2 treatment groups was similar in terms of both the number of drop-outs at any stage and the reason for discontinuation (see Tables 9 and 10).

Table 9: Disposition of subjects by group (Study #0003)

	Zoladex: 4-week		Zoladex: 8-week	
	n	rand. %	n	rand. %
Randomized	40	100.0	40	100.0
Received treatment	40	100.0	38	95.0
Completed treatment	39	97.5	38	95.0
Received surgery	38	95.0	38	95.0
Completed trial	37	92.5	37	92.5

Source: Vol. 18.28, Tables T3, T4, T5

Table 10: Reasons for Discontinuation (Study #0003)

	Zoladex: 4-week		Zoladex: 8-week	
	n	% rand.	n	% rand.
Withdrew before treatment				
Patient request	0	0.0	1	2.5
Other	0	0.0	1	2.5
Withdrew during trmt; before surgery				
Patient request	1	2.5	0	0.0
Withdrew during or after surgery				
Adverse event (prior to ELA)	1	2.5	0	0.0
Did not attend post-surgery follow-up visit	1	2.5	1	2.5

Source: Vol. 18.28, Tables T3, T4, T5

Applicant's Analysis

The study report for study #0003 did not specify the subject population to be included in the applicant's analyses. However, by checking group counts in the data, this reviewer determined that the efficacy analyses reported by the applicant were conducted on the same modified intent-to-treat (ITT) basis as in study #0022. The true ITT (all randomized subjects) was 'modified' in the sense that patients with no data for an endpoint were excluded in the analysis of that endpoint. All randomized subjects who had data for an endpoint were included in the analysis of that endpoint. Two (2) randomized subjects from each treatment group did not have surgery (thus no variables measured regarding surgery). An additional 1 randomized subject from each treatment group did not compete the 24-week follow-up visit so no information was recorded for the amenorrhea response variable. See Table 9 on previous page.

The efficacy variables were measured on different scales which required different methods of analysis. For each variable, the measurement scale, applicant's analysis method, and applicant's results, are described below.

Incidence of Amenorrhea (primary efficacy variable)

The incidence of amenorrhea at 24 weeks post-surgery is a binary variable. In this study, a categorical response variable, improvement in menorrhagia, was recorded, with amenorrhea as one of the 4 categories. The applicant collapsed this to 2 categories: Amenorrhea, and Non-amenorrhea. A Fisher's Exact test was performed on the 2x2 table to test for treatment differences between the 4-week and 8-week Zoladex regimens. The amenorrhea rates were the same for the 2 groups and there was no treatment difference found.

Table 11a: Applicant's Results (Modified ITT) (Study #0022)

	Zoladex 4-week (n=37)		Zoladex 8-week (n=37)		Fisher's Exact Test Results
	n	%	n	%	
Incidence of Amenorrhea (%)	9	24.3	9	24.3	(p-value=1.00)

Source: Vol. 18.28, Table T22.2

Endometrial Thickness (primary efficacy variable)

Endometrial thickness immediately pre-surgery was measured on a continuous scale (mm). The data was not normally distributed, so the applicant used a non-parametric test, the Wilcoxon Rank-Sum test, to compare the Zoladex 4-week and Zoladex 8-week treatment groups. (The study report refers to the results as being from the Mann-Whitney U-test, but the analysis performed was the Wilcoxon Rank-Sum test.) The Wilcoxon procedure was also used to generate a 95% confidence interval on the median difference.

Note that the arithmetic difference between the medians for the 2 groups does not necessarily equal the median of the differences estimated by the Wilcoxon method. This is because the Wilcoxon method calculates the difference between all possible pairs of subjects ($[n_1 \times n_2] \div 2$ = total # of pairs) and then calculates the median of all differences, not the difference of the 2 group medians.

The results indicate a significant treatment effect between the Zoladex 4-week and Zoladex 8-week treatment groups. The direction of the difference favors the Zoladex 8-week regimen, as a positive value for the difference relates to smaller endometrial thickness in the Zoladex 8-week treatment group.

Table11b: Applicant's Results (Modified ITT) (Study #0022)

	Zoladex 4-week		Zoladex 8-week		Mann-Whitney U-test Results	
	n	Median (25th, 75th p-tile)	n	Median (25th, 75th p-tile)	Median Difference	95% Confidence Interval
Endometrial Thickness (mm)	39	1.00 (1.00, 1.00)	38	0.50 (0.50, 0.70)	0.50 (p-value=.0001)	(0.0, 0.5)

Source: Vol. 18.28, Table T8.2.

Duration of Surgery (secondary efficacy variable)

Duration of surgery was measured on a continuous scales (minutes), and the data fit the assumptions for an ANCOVA model. The model used included a factor for treatment and baseline uterine volume as the covariate. The results of this model are shown in Table 11c below. There was no significant treatment effect, after adjusting for baseline uterine volume.

Table 11c: Applicant's Results (Modified ITT) (Study #0003)

	Zoladex 4-week		Zoladex 8-week		ANCOVA Model Results	
	n	Adjusted * mean (Std. Dev)	n	Adjusted * mean (Std. Dev)	LS Means Difference	95% Confidence Interval
Duration of Surgery (min)	38	16.91 (4.45)	38	15.30 (4.45)	1.61 (p-value=.1196)	(-0.43, 3.65)

Source: Vol. 18.28, Table T17.2.
* Adjusted for baseline uterine volume in the ANCOVA model

Difficulty of Surgery (secondary efficacy variable)

Ease of surgery was measured on a 4-point ordinal scale. The categories were: None; Minimal; Moderate; and Severe. The parameters contributing to difficulty were not specified at the time of surgery, however these generally included factors such as difficulty in dilation of the cervix, poor endometrial preparation, or poor visibility at surgery (Vol. 18.28, pg. 40). The results are given in Table 11d. No statistical testing was done for this variable by the applicant, and no conclusions were reported.

Table 11d: Applicant's Results (Modified ITT) (Study #0003)

	Zoladex 4-week		Zoladex 8-week	
	n	proportion	n	proportion
Difficulty of Surgery	38	100.0	38	100.0
None	27	71.1	26	68.4
Minimal	9	23.7	9	23.7
Moderate	1	2.6	3	7.9
Severe	1	2.6	0	0.0

Source: Vol. 18.28, Table T18.2.

Reviewer's Analysis

The modified intent-to-treat population of subjects which the applicant used includes all randomized subjects who had data for each variable. This same population is used for this reviewer's analyses.

Incidence of Amenorrhea (primary efficacy variable)

The protocol for study #0003 (Vol. 18.28, pg. 173) did not mention the amenorrhea variable as an efficacy variable, nor did it discuss the analysis of this variable. For consistency with the analysis of study #0022, this reviewer calculated the difference and 95% confidence intervals for the amenorrhea rates. These appear in Table 12a below. As in the applicant's analysis, there is no significant treatment difference between the Zoladex 4-week and Zoladex 8-week regimens.

Table 12a: Reviewer's Results (Modified ITT) (Study #0003)

	Zoladex 4-week (n=37)		Zoladex 8-week (n=37)		Difference in Rates	
	n	%	n	%	Difference	95% Confidence Interval
Incidence of Amenorrhea (%)	9	24.3	9	24.3	0.0	(-19.6, 19.6)

Endometrial Thickness (primary efficacy variable)

In the protocol (Vol. 18.28, pg. 174) the proposed analysis for endometrial thickness is a 2-sample t-test comparing the Zoladex 4-week and Zoladex 8-week treatment groups. If the assumptions for this parametric test were not met, then a non-parametric analysis with ranks was planned. The Wilcoxon Rank-Sum test reported in the study analysis is a non-parametric test based on ranks and is appropriate for this variable. No further analysis was done by this reviewer.

Duration of Surgery (secondary efficacy variable)

In the protocol (Vol 18.28, pg. 174), the proposed model for the analysis of duration of surgery was a 2-sample t-test. This is equivalent to an ANOVA model with only a factor for the 2 treatment groups. The applicant performed an ANCOVA model with a factor for the 2 treatment groups plus the covariate term for baseline uterine volume. This reviewer calculated the 2-sample t-test for treatment effect for comparison to the applicant's ANCOVA results. No significant treatment effect was found from the t-test results, which is consistent with the applicant's results.

Table 12b: Reviewer's Results (Modified ITT) (Study #0003)

	Zoladex 4-week		Zoladex 8-week		Two Sample T-test Results	
	n	Mean (Std. Dev.)	n	Mean (Std. Dev.)	Difference in Means	95% Confidence Interval
Duration of Surgery (min)	38	17.00 (4.44)	38	15.21 (4.67)	1.79 (p-value=.0911)	(-0.26, 3.84)

Difficulty of Surgery (secondary efficacy variable)

The protocol for study #0003 proposed analyzing this variable using a Mantel-Haenszel χ^2 test. This reviewer performed this test, and found no significant treatment effect on difficulty of surgery.

Table 12c: Reviewer's Results (Modified ITT) (Study #0003)

	Zoladex 4-week		Zoladex 8-week		Mantel-Haenszel χ^2 Results
	n	proportion	n	proportion	
Difficulty of Surgery	38	100.0	38	100.0	p-value = 0.860
None	27	71.1	26	68.4	
Minimal	9	23.7	9	23.7	
Moderate	1	2.6	3	7.9	
Severe	1	2.6	0	0.0	

Source: Vol. 18.28, Table T18.2.

Subgroup Analyses

For descriptive purposes only, this reviewer compared the primary efficacy results for each treatment group by age subgroups. This is a single-center study, so no analysis by center was needed.

For the age comparison, two age groups were defined by this reviewer. The mean age in the treatment groups was 42.1 for the Zoladex 4-weeks group and 41.2 for the Zoladex 8-weeks group, so the age groups were split at: 41 and under; and 42 and older to be consistent with the subgroups analysis for study #0022.

Within each age group, the 2 treatment regimens had similar results.

Table 13a: Age Group Comparison (Modified ITT) (Study #0003)

	Age 41 and Under				Age 42 and Older			
	Zoladex 4-weeks (n=14)		Zoladex 8-weeks (n=19)		Zoladex 4-weeks (n=21)		Zoladex 8-weeks (n=18)	
	n	%	n	%	n	%	n	%
Incidence of Amenorrhea (%)	3	21.4	4	21.1	6	28.6	5	27.8

For endometrial thickness immediately pre-surgery, the trend across age group was consistent with the treatment difference found in the main analysis (Table 11b). Specifically, subjects in the Zoladex 8-week treatment group had a lower endometrial thickness in both age groups than the subjects in the Zoladex 4-week treatment group.

Table 13b: Age Group Comparison (Modified ITT) (Study #0003)

	Age 41 and Under		Age 42 and Older	
	Zoladex 4-weeks (n=19)	Zoladex 8-weeks (n=20)	Zoladex 4-weeks (n=20)	Zoladex 8-weeks (n=18)
	Unadjusted Median	Unadjusted Median	Unadjusted Median	Unadjusted Median
Endometrial Thickness (mm)	1.0	0.5	1.0	0.5

At the request of the medical officers, this reviewer considered additional exploratory subgroup comparisons for baseline characteristics for study #0003. The goal was to determine if there were advantages for either the Zoladex 4-week or Zoladex 8-week treatment regimen based on subject characteristics at baseline. Descriptive analysis techniques were used to investigate possible trends in treatment effect for the 2 primary efficacy variable, incidence of amenorrhea and endometrial thickness immediately pre-surgery, for each of the following baseline characteristics:

Uterine size (weeks)
Uterine volume (cc)
Uterine cavity length
Baseline endometrial thickness
Presence of submucous fibroids

The main analysis of incidence of amenorrhea (Tables 11a and 12a) concluded there was no significant treatment effect. The same conclusion was indicated for all of the baseline characteristics. For endometrial thickness immediately pre-surgery, the full analysis (Table 11b) found a significant treatment effect favoring the Zoladex 8-week treatment regimen. The subgroup analysis found this trend favoring the Zoladex 8-week treatment regimen was consistent across each of the 5 baseline variables, but statistical significance was not tested. It is important to note that this subgroup analysis is only for descriptive purposes and is based on rather small sample sizes. It is not possible to make any definite conclusions based on this descriptive analysis.

Conclusions - Study #0003

Study #0003 compared 2 regimens for the Zoladex treatment: a 4-week (1 injection) treatment, and an 8-week (2 injection) treatment. This is a subset of the 4 treatment groups included in the full clinical trial. Two other groups received Danazol, an unapproved treatment. In discussions with the FDA prior to submission of this NDA (pre-NDA mtg. Oct. 26, 1995), it was agreed that Danazol was considered to be investigational for this indication in the U.S. and that the 2 Danazol groups would not be considered to be suitable positive control comparators for this submission.

The primary efficacy variables for this study were incidence of amenorrhea and endometrial thickness immediately pre-surgery. The results indicate a significant treatment effect for endometrial thickness (p-value=.0001), in a direction which favored the 8-week regimen as having a smaller median thickness immediately pre-surgery. However, no significant treatment effect was found between the 2 Zoladex regimens for the incidence of amenorrhea (p-value=1.00). There were 2 secondary efficacy variables in the analysis: duration of surgery and difficulty of surgery. No significant treatment effects were found for either of these variables.

The results of study #0003 are reported in the proposed label only for the 2 Zoladex treatment groups combined, not as the 2 separate regimens considered in the efficacy analysis. The results for endometrial thickness are not reported, even though this was the only efficacy variable for which a treatment effect between the 2 regimens was found. Also, the results for amenorrhea are reported in the label using categories different from those in the analysis. The variable which was recorded during the collection of data was Improvement in menorrhagia, a 4-category variable coded as Amenorrhea, Hypomenorrhea, Normal Flow, or Not Improved. For the efficacy analysis, this was collapsed as Incidence of amenorrhea, which is a binary variable (Vol. 18.28, pg. 90). The label should present the results as analyzed and reported in the study report.

Summary (0022 & 0003)

The two clinical trials in this NDA assessed the efficacy of Zoladex for endometrial thinning prior to endometrial ablation in different ways. Study #0022 compared Zoladex (2 injections) to a sham treatment (2 injections), with surgery 6 weeks after the first injection (2 weeks after 2nd injection). The endometrial ablation procedure was by loop resection. Study #0003 compared 2 Zoladex treatment regimens (1 injection vs. 2 injections) with surgery 4 weeks after the last injection, and used the endometrial laser ablation procedure.

Table 14: Summary of Efficacy Results

	Study #0022 Zoladex vs. Sham			Study #0003 Zoladex 4-week regimen vs. Zoladex 8-week regimen		
Variable	Statistic Used (All diff. are Zol. - Sham)	Point Estimate (* stat. sign.)	95% Confidence Interval	Statistic Used (All diff. are 4 wk - 8-wk)	Point Estimate (* stat. sign.)	95% Confidence Interval
Incidence of Amenorrhea (%)	Difference in %	14.3 *	(4.5, 24.0)	Difference in %	0.0	(-19.6, 19.6)
Endometrial Thickness (mm)	Median Difference	-1.85 *	(-2.2, -1.5)	Median Difference	0.5 *	(0.0, 0.5)
Change in Blood Loss Score	Difference in Means	-14.6	(-96.6, 67.4)	Not measured in study 0003		
Duration of Surgery (minutes)	Median Difference	-4 *	(-6, -2)	Difference in Means	1.79	(-0.26, 3.84)
Ease of Surgery	Odds Ratio	6.45 *	(4.14, 10.04)	Not measured in study 0003		
Difficulty of Surgery	Not measured in study 0022			Chi-square test * stat. sign.	not applicable	not applicable

Source: Tables 5e, 6a, 6b, 6c, 6e, 11b, 12a, 12b.

Two of the primary efficacy variables were considered in both studies. For endometrial thickness immediately prior to surgery, a significant difference between Zoladex (2 injections) and the sham treatment was found in study #0022, and the results favored Zoladex, with the Zoladex treatment group having a smaller median endometrial thickness immediately before surgery. The point estimate for the median difference was -1.85 mm, with a 95% confidence interval of (-2.2, -1.5). In study #0003, there was a significant difference between the Zoladex 4-week (1 injection) and 8-week (2 injections) regimens which favored the 8-week regimen as having smaller median endometrial thickness. The median difference was 0.5 mm, with a 95% confidence interval of (0.0, 0.5). The estimated median thickness (mm) for the Zoladex groups was 1.5 mm for Zoladex (2-injections) in study #0022; 1.0 for Zoladex (1-injection) in study #0003; and 0.5 for Zoladex (2-injections) in study #0003. Note that the study plans differed in when the surgery was scheduled after the last injection: 2 weeks in study #0022, 4 weeks in study #0003.

The other primary efficacy variable which the 2 studies had in common was the incidence of amenorrhea at 24 weeks post-surgery. The results from study #0022 indicate a significant treatment difference which favored Zoladex (2 injections) over the sham treatment. However, the 95% confidence interval on the difference, (4.5, 24.0), included the value of 20% which was defined in the protocol as clinically meaningful. Thus the results indicate a significant difference, but it may not be a clinically meaningful difference. Study #0003 showed no

significant difference in the incidence rate between the 2 Zoladex regimens. The difference between the 2 regimens was 0.0%, with a 95% confidence interval of (-19.6, 19.6). The estimated incidence rate for amenorrhea in the Zoladex (2 injections) group in study #0022 was 40.0% (superior to sham: 25.7%), but in study #0003 the incidence rate for both Zoladex treatment groups was 24.3%.

In study #0022, a third primary efficacy variable, change in blood loss score (baseline to 24 weeks post-surgery), was specified by the applicant. In discussions with the medical officers, it was decided that this variable was not of primary concern for the regulatory decision. The results indicated no significant difference between the Zoladex and sham treatment groups. This contradicts the results of the incidence of amenorrhea response variable, yet further investigation provided no valuable clarification of this contradiction.

Secondary efficacy variables were related to the surgery. Duration of surgery was found to be significantly shorter for the Zoladex (2 injections) treatment than for the sham treatment in study #0022, with no differences between the 2 Zoladex groups found in study #0003. The ease of surgery was also significantly different for the Zoladex (2 injections) treatment than for the sham treatment in study #0022. The estimated odds ratio of having easier surgery was 6.45, indicating subjects receiving Zoladex were more than 6 times more likely to have easier than normal surgery than subjects in the sham group. In study #0003, difficulty of surgery was measured instead of ease of surgery (not on comparable scales). There was no difference between the 2 Zoladex regimens for difficulty of surgery.

The following comments relate to the proposed label (Vol. 18.1, pg. 34-36):

Katherine B. Meaker

Katherine B Meaker, M.S.
Mathematical Statistician

Concur: Dr. Nevius 6/15/97
Dr. Kammerman 6/2/97

cc:
Archival NDA 19-726 / S018
HFD-580
HFD-580/CCropp, HJolson, LRarick
HFD-580/ADunson
HFD-715/ENevius, LKammerman, KMeaker, Chron

KBM/06/02/97 \Word 7.0\My Documents\NDA19726-zoladex\Review