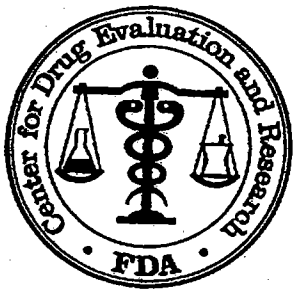
CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

22-045

STATISTICAL REVIEW(S)



US Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

Statistical Review and Evaluation

New Drug Application

Clinical Studies

NDA/Serial Number:	22-045/SN000
Drug Name:	YAZ™
Indication(s):	Moderate Acne
Applicant:	Berlex, Inc.
Dates:	Submitted: March 27, 2006 PDUFA: January 26, 2007
Review Priority:	Standard Review
Biometrics Division:	Division of Biometrics III
Statistics Reviewer:	Clara Kim, Ph.D.
Concurring Reviewer:	Mohamed Alosh, Ph.D.
Medical Division:	Division of Dermatology and Dental Products
Clinical Team:	David Kettl, M.D./Markham Luke, M.D.
Project Manager:	Mildred Wright
Keywords:	moderate acne, oral contraceptive

Contents

1 Executive Summary

1.1 Conclusions and Recommendations

YAZ tablets has been demonstrated to be statistically superior to placebo in two studies (Study 306820 and Study 309669) in the treatment of moderate acne. Efficacy was evaluated on (i) success rate based on the Investigator Static Global Assessment (ISGA) score and (ii) percent change from baseline for two out of three lesions counts (inflammatory, non-inflammatory, and total) at Day 15 of Treatment Cycle 6. Table 1 presents the summary of the co-primary endpoint results. All co-primary endpoints were statistically significant in both studies with p-values less than 0.0003.

Table 1: Efficacy Results Summary - Number (Proportion) of success on ISGA and Mean Baseline and % Reduction (SD) in Lesion Counts

	Study 306820		Study 306996	
	YAZ N=228	Placebo N=230	YAZ N=218	Placebo N=213
ISGA				
Number of Successes	35 (15.4%)	10 (4.3%)	46 (21.1%)	19 (8.9%)
p-value*		0.0001		0.0003
Inflammatory				
Baseline count	32.61 (16.1)	32.84 (14.6)	31.73 (12.4)	31.81 (13.7)
% reduction	47.60% (35.4%)	32.34% (37.3%)	50.60% (38.3%)	34.46 (49.7%)
p-value†		<0.0001		<0.0001
Non-inflammatory				
Baseline count	47.27 (31.4)	46.99 (30.6)	43.85 (22.9)	43.85 (25.9)
% reduction	38.08% (39.0%)	18.22% (47.9%)	42.35% (38.5%)	26.02% (48.2%)
p-value†		<0.0001		<0.0001
Total				
Baseline count	79.88 (42.3)	79.83 (37.2)	75.60 (30.7)	75.67 (33.9)
% reduction	42.33% (32.7%)	25.29% (36.4%)	46.13% (33.7%)	30.64% (41.9%)
p-value†		<0.0001		<0.0001

* p-values are calculated using a logistic model with treatment and pooled centers as factors

† p-values are calculated using ANCOVA model: Baseline lesion count as covariate, and treatment, center, and treatment by center interaction (if statistically significant) as factors.

Source: Reviewer analysis

The adverse event rates were higher in YAZ subjects than in the placebo subjects. The most common adverse event was upper respiratory tract infection, which was reported by approxi-

mately 9% of the subjects. The next common adverse event was metrorrhagia.

1.2 Brief Overview of Clinical Studies

The sponsor conducted two Phase 3 studies (Study 306820 and Study 306996) evaluating the safety and efficacy of YAZ tablets versus its placebo in the treatment of moderate acne. A total of 534 and 538 subjects with acne were randomized in a 1:1 ratio to either YAZ or placebo from Studies 306820 and 306996, respectively. Subjects were on treatment for 6 treatment cycles (168 days). Efficacy was evaluated on Day 15 of Treatment Cycle 6 for the following primary endpoints: (i) proportion of subjects who had a score of 0 or 1 based on the ISGA; (ii) percent change from baseline for two out of three lesion counts (inflammatory, non-inflammatory, or total). All study centers were in the United States.

1.3 Statistical Issues and Findings

The sponsor conducted two studies (Study 306820 and Study 306996) under the protocol that was agreed upon with the Agency in terms of study design and endpoints. Efficacy was evaluated on Day 15 of Treatment Cycle 6 for the proportion of success rate based on the ISGA score and the percent change in lesion (inflammatory, non-inflammatory, and total) counts from baseline. The difference in the success rates were statistically significant in both studies (p-values <0.0003). The difference in the mean percent changes in lesion counts were also statistically significant in all lesion types in both studies (p-values <0.0001). Within each study, the efficacy results were relatively consistent across subgroups and investigative sites. However, in Study 306820, Site 6 (200188=[redacted]) showed larger treatment effects compared to that of other study sites. Sensitivity analyses which excluded this extreme center still demonstrated superiority of YAZ tablets to placebo. The protocol defined seven secondary endpoints, which had little regulatory utility or were not relevant to labeling claims.

2 Introduction

2.1 Overview

YAZ™ tablets (Drospirenone 3mg and Ethinyl Estradiol-β-Cyclodextrin Clathrate 0.02mg) was approved on March 16, 2006 by the Division of Reproductive/Urologic Drug Products (DRUP) for the indication of prevention of pregnancy (NDA 21-676).

Three European studies evaluated the efficacy and safety of the combination of Drospirenone (DRSP) 3mg and Ethyinyl Estradiol (EE) 0.03mg (Yasmin®) for the treatment of moderate acne vulgaris, seborrhea, and hirsutism. The sponsor stated that all 3 studies showed favorable efficacy and safety results for the treatment of moderate acne vulgaris, and that the out-

comes of these studies were the basis of evaluating DRSP 3mg in combination with EE 0.02mg (herein referred to as DRSP/EE or YAZ) for the treatment acne. In the current application of YAZ™ tablets, the sponsor is seeking an indication of moderate acne vulgaris as an efficacy supplement, in otherwise healthy women of reproductive age of 14 to 45 years, treated for 6 treatment cycles. One treatment cycle consists of 1 tablet containing active substance a day for 24 days followed by a 4-day inert tablet, or 28 placebo tablets.

Two identical Phase 3 study protocols were submitted as Special Protocol Assessment (SPA) on October 28, 2002. Comments were conveyed to the sponsor on December 12, 2002. The sponsor requested a type A meeting to clarify some comments, which was held on February 4, 2003. The Agency requested that oral contraceptives are not indicated for ' —— ' acne and that the indication should be changed to "moderate acne vulgaris" from " ———— ". Consequently, the sponsor amended the inclusion criteria in the early course of the study. The original inclusion criteria was 10 to 100 non-inflammatory (comedones), 10 to 50 inflammatory (papules or pustules) lesions, and not more than 5 nodules with an ISGA score of 2 and above. This criteria was changed to a minimum of 40 lesions with at least 20 inflammatory and 20 non-inflammatory lesions with not more than 3 nodules on the face with an ISGA score of 3 and above. The Agency also requested the sponsor to change the ISGA scale from a 5-grade scale to a 6-grade scale. There were 4 subjects in Study 306820 and 6 subjects in Study 306996 that were included in the sponsor's efficacy analyses with an ISGA score of 2 measured using the 5-grade ISGA scale prior to the Agency's request. The sponsor's efficacy analysis of Study 306820 excluded 6 subjects who had baseline inflammatory lesion counts of 20 or 21. In this review, the population that excludes all subjects with baseline ISGA score of 2, regardless of the scale used, and includes the 6 subjects with baseline inflammatory lesion counts of 20 or 21 will be denoted as the 'modified' population, on which the primary analysis is based. The sponsor's efficacy analysis population, based on the modified inclusion criteria, will be denoted as the 'amended intent-to-treat (ITT)' population.

Through the SPA review and type A meeting, the sponsor and Division came to an agreement on endpoints and most aspects of the study design. Table 2 lists the clinical study programs and the number of subjects. This review evaluates the two Phase 3 studies' efficacy based on the modified and amended ITT populations. Safety is evaluated on the sponsor's original ITT population, which does not reflect the change in inclusion criteria (534 and 538 subjects in Studies 306820 and 306996, respectively).

Table 2: Overview of Pivotal Clinical Studies

Study	Study Period	Enrollment (Modified Population)			Enrollment (Amended ITT)		
		YAZ	Placebo	Total	YAZ	Placebo	Total
306820	1/13/03 – 7/15/04	228	230	458	229	227	456
306996	1/23/03 – 6/22/04	218	213	431	222	215	437

2.2 Data Sources

This reviewer evaluated the sponsor's clinical study reports and clinical summaries, as well as the proposed labeling. This submission was submitted in CTD format and was entirely electronic. The datasets used in this review are archived at \\Cdsesub1\n22045\N_000\2006-03-24\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\acne-vulgaris\5351-stud-rep-contr\study-report-306820\datasets\tabulations and \\Cdsesub1\n22045\N_000\2006-03-24\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\acne-vulgaris\5351-stud-rep-contr\study-report-306996\datasets\tabulations.

3 Statistical Evaluation

3.1 Evaluation of Efficacy

3.1.1 Study Design

The sponsor conducted two Phase 3 studies (Study 306820 and Study 306996) to evaluate the safety and efficacy of YAZ™ tablets in the treatment of moderate acne vulgaris as an efficacy supplement. Studies 306820 and 306996 were conducted under identical protocols, which were submitted as SPA on October 28, 2002. Comments were conveyed to the sponsor on December 12, 2002. The Agency requested that the indication should be limited to "moderate" acne [redacted] acne vulgaris. The sponsor changed the inclusion criteria according to the Agency's request in their protocol amendment 3, dated February 3, 2003. This was shortly after the sponsor started their studies (1/13/03 and 1/23/03 for Studies 306820 and 306996, respectively). The sponsor continued to enroll subjects according to the original inclusion criteria, however excluded subjects who did not meet the amended inclusion criteria from the efficacy analyses. This population is denoted as 'amended ITT population' in this review, hereafter. (The sponsor denoted this population as 'amended full analysis set'.) The sponsor also changed the ISGA scale from a 5-grade scale to a 6-grade scale at the Agency's request at this time. The subjects enrolled in the studies were females between 14 and 45 years of age, $\geq$ 1 year post-menarche with a minimum of 40 lesions with at least 20 inflammatory

lesions (papules or pustules), 20 non-inflammatory (comedones), not more than 3 small inactive nodules and not classified as Grade 0, 1, or 2 on the ISGA scale. (The definition of the ISGA scale is listed below.)

The studies were designed as multicenter, randomized, double-blind, placebo-controlled studies. Each study planned to enroll a total of approximately 500 subjects from up to 50 sites. The actual enrollment was 534 and 538 subjects, from 29 and 23 study centers for Studies 306820 and 306996, respectively. The actual numbers of subjects included in the amended ITT population for studies 306820 and 306996 were 456 and 437, from 24 and 22 study centers, respectively. The excluded subjects (78 and 101 subjects from Studies 306820 and 306996, respectively) had a baseline ISGA score of 0-2 on the 6-grade scale, or had a baseline inflammatory or non-inflammatory lesion count less than 20. Note that subjects who entered the study before the ISGA scale was changed and had a ISGA score of 2 were included in the amended ITT population. In this review, 'modified population' was defined as the population that excluded all subjects with a baseline ISGA score of 2, regardless of the scale used, and included subjects with baseline inflammatory lesion count of 20 or 21, who were excluded from the amended ITT population.

The enrolled subjects were randomly assigned to YAZ or placebo in a 1:1 ratio. In the modified population, a total of 458 and 431 subjects met the inclusion criteria from Studies 306820 and 306996, where 228 and 230 subjects were randomized to YAZ and placebo in the first study and 218 and 213 subjects were randomized to those arms in the second study. In the amended ITT population, a total of 456 and 437 subjects were included in the efficacy analysis in Studies 306820 and 306996, where 229 and 227 subjects were randomized to YAZ and placebo in the first study and 222 and 215 subjects were randomized to those arms in the second study. The following treatments were administered for 6 treatment cycles:

YAZ: DRSP/EE oral tablet daily for 24 days, oral inert tablet for 4 days.

Placebo: Look-alike DRSP/EE placebo oral tablet daily for 24 days, oral inert tablet daily for 4 days.

The initial (Baseline) facial lesion count and ISGA was assessed 15±3 days after the first day of menstrual bleeding. Subjects were instructed to take the first tablet on the first day of their menstrual cycle. Additional efficacy evaluations were conducted on Day 15±3 of Treatment Cycles 1, 3, and 6. Treatment Cycle 6 was the primary time point for efficacy assessment.

The following endpoints were specified in the protocol for efficacy evaluation.

- Primary endpoints (evaluated at Treatment Cycle 6):
  - Percentage of subjects classified as 0 or 1 on the ISGA scale
  - Percent change from Baseline in inflammatory lesion count
  - Percent change from Baseline in non-inflammatory lesion count

  - Percent change from Baseline in total lesion count

- Secondary endpoints:

  - Change from Baseline in individual lesion counts (open comedones, closed comedones, papules, pustules, nodules)
  - Percentage of subjects classified as "improved" according to the Investigator's Overall Improvement Rating
  - Percentage of subjects classifying themselves as "improved" on the Subject's Overall Self-Assessment Rating

Inflammatory lesions included papules, pustules, and nodules, whereas non-inflammatory lesion included open and closed comedones. The ISGA was based on a 6-grade scale in response to the Agency's clinical comments conveyed to the sponsor in December, 2002. The following remarks were also conveyed to the sponsor in the same set of comments regarding the secondary endpoints. The statistical reviewer stated "a multiplicity adjustment would be needed if the number of secondary efficacy endpoints intended for labeling is large". The clinical comments included "The proposed dynamic assessments by investigator and subject [Investigator's Overall Improvement Rating and Subject's Overall Self-Assessment Rating] have little regulatory utility. The Agency prefers to rely, when available as is the case for acne, on more objective measurements such as lesion counts and the static assessments at the time of evaluation rather than on 'subjective impressions' of degree of improvement which may rely on less than perfect recollection of previous status." The 6-point ISGA scale is defined as the following.

Score	Definition
0	Normal, clear skin with no evidence of acne vulgaris
1	Skin is almost clear: few non-inflammatory lesions present, with rare non-inflamed papules (papules must be resolving and may be hyperpigmented, though not pink-red), no nodular lesions
2	Few inflammatory lesions (papules or pustules), little inflammation, some comedones, no nodular lesions
3	Many comedones (non-inflammatory lesions predominate), several inflammatory lesions (papules or pustules), one small nodular lesion may or may not be present
4	Many inflammatory lesions (papules and pustules), up to many comedones, there may not be a few nodular lesions
5	Numerous highly inflammatory lesions predominate. Variable number of comedones, many papules and pustules or nodular lesions

The Investigator's Overall Improvement Rating and Subject's Overall Self-Assessment Rating were also on 6-point scales, defined as the following.

1	Clear	Complete clearance of disease signs and symptoms compared to Baseline
2	Excellent improvement	> 75% clearance of disease signs and symptoms compared to Baseline
3	Good improvement	50% to 75% clearance of disease signs and symptoms compared to Baseline
4	Moderate improvement	< 50% clearance of disease signs and symptoms compared to Baseline
5	No improvement	No appreciable clearance of disease signs and symptoms compared to Baseline
6	Deterioration	Worsening of disease signs and symptoms compared to Baseline

The submission defined the ITT population as all randomized subjects who were dispensed trial medication and stated that the protocol incorrectly defined the ITT population to exclude subjects to whom the study medication was not administered. This reviewer used the submission's definition of the ITT population. In the submission, the sponsor treated data collected after Day 168 as missing. This visit window was added to the Statistical Analysis Plan (SAP) dated 11/14/03. This addition was not reviewed by the Agency. The sponsor stated that the SAP was signed off before the database was locked. The modified population is based on the data set that incorporates these visit windows. The amended ITT population utilizes all data observed as defined in the protocol and disregards the visit windows. In this review, the primary analysis will be based on the modified population to be consistent with the sponsor. Supportive analyses will be based on the amended ITT and per protocol populations. The sponsor's analysis of the primary endpoints, which were based on the amended ITT population, but utilized the visit windows are provided as a sensitivity analysis.

The analysis methods proposed in the protocol are the following. Unless stated otherwise, the proposed analysis methods were used in the submission.

- The sponsor proposed to analyze the success rate (ISGA) at Day 15±3 days of Cycle 6 using a logistic regression model with treatment and center as factors in the original protocol. The Division recommended to include the treatment by center interaction in the model. In their amended protocol, they proposed to investigate the interaction term at the 0.10 significance level as an exploratory analysis. In this review, the interaction term is included in the analysis model unless it is determined not statistically significant at $\alpha = 0.10$.

- Lesion counts were analyzed using an Analysis of Covariance (ANCOVA) model with percent change from Baseline to Day 15±3 days of Cycle 6 as the response variable and baseline value as a covariate. Similarly to ISGA, the sponsor proposed to include treatment and center as factors in the ANCOVA model. In response to the Division's recommendation to include the treatment by center interaction term in the analysis model, the sponsor amended the protocol to investigate the interaction term at a significance level of 0.10. In this review, the interaction term is included in the ANCOVA model unless it is determined not statistically significant at $\alpha = 0.10$.

- Secondary endpoints, change from Baseline in individual lesions counts, was analyzed at

Day 15±3 days of Cycle 6 using the same ANCOVA model described as the primary endpoint (inflammatory and non-inflammatory lesion counts).

- Secondary endpoints, Investigator's Overall Improvement Rating and Subject's Global Assessment Scale, was analyzed at Day 15±3 days of Cycle 6 in the same manner as primary endpoint, ISGA.
- Missing data were imputed using Last-Observation-Carried-Forward (LOCF) approach. Sensitivity analyses regarding the imputation method of missing data were not proposed in the protocol, nor conducted in the submission. The reviewer's sensitivity analysis is provided in section 3.1.4.2.
- The target number of subjects per site per arm was at least 10 subjects. Sites that did not meet the target were ranked with respect to the total number of subjects. Small sites were pooled from the largest to the smallest until the pooled center had at least 10 subjects per arm per center. Study sites were pooled into 11 investigational groups in both studies.

3.1.2 Subject Disposition

Based on the modified population, Study 306820 enrolled 458 subjects who met the inclusion criteria and randomized 228 to YAZ and 230 to placebo, at 24 study centers. Study 306996 enrolled 431 subjects who met the inclusion criteria and randomized 218 to YAZ and 213 to placebo, at 22 study centers. Table 3 presents the reasons for study discontinuations.

Table 3: Number (Proportion) of Subjects Who Discontinued the Study: Classified by the Reason for Discontinuation

	Study 306820		Study 306996	
	YAZ N=228	Placebo N=230	YAZ N=218	Placebo N=213
Subjects who discontinued	**61 (27%)**	**65 (28%)**	**48 (22%)**	**62 (29%)**
Reason				
Withdrawal of Consent	14 (6%)	12 (5%)	11 (5%)	15 (7%)
Protocol Deviation	5 (2%)	2 (<1%)	0 (0%)	4 (2%)
Adverse Event	15 (7%)	9 (4%)	14 (6%)	7 (3%)
Patient Loss	13 (6%)	20 (9%)	16 (7%)	18 (8%)
Pregnancy	3 (1%)	2 (<1%)	4 (2%)	7 (3%)
Lack of Efficacy	0 (0%)	3 (1%)	0 (0%)	2(<1%)
Other	11 (5%)	16 (7%)	3 (1%)	8 (4%)

Source: Reviewer's analysis

The number of subjects who discontinued the study was slightly higher in the placebo arm than the YAZ arm in both studies. The most common reasons for study discontinuation in YAZ subjects were adverse events and patient loss, whereas in the placebo arm, patient loss and withdrawal of consent were most common. The proportions of subjects classified as 'other' were above 5% in Study 306820 for both arms. The most common reasons for discontinuation categorized as 'other' in the YAZ arm were due to the enrollment timeline (5 subjects) and subjects being non-compliant with study visits (4 subjects). In the placebo arm, 10 subjects of the 17 'other's discontinued due to non-compliance with the study medication.

3.1.3 Baseline and Demographic Data

Table 4 presents the baseline demographic data based on the modified population. Baseline demographic variables were generally balanced across treatment arms. The average age of all subjects was 24.5 years and the age range was from 14 to 45 years. For both studies, more Caucasians were randomized to YAZ than placebo. The discrepancy in the proportions of Caucasians of the two arms were 8% and 5% for Studies 306820 and 306996, respectively.

Table 4: Baseline Demographics

	Study 306820		Study 306996	
	YAZ N=228	Placebo N=230	YAZ N=218	Placebo N=213
Age (in years)				
mean (std)	24.5 (6.7)	25.1 (6.6)	24.7 (7.6)	24.6 (7.5)
median	23	25	23.5	23
min, max	(14,44)	(14,45)	(14,44)	(14,43)
Race				
Caucasian	161 (71%)	146 (63%)	157 (72%)	142 (67%)
African-American	31 (14%)	41 (18%)	34 (16%)	27 (13%)
Hispanic	23 (10%)	26 (11%)	20 (9%)	28 (13%)
Asian	4 (2%)	7 (3%)	5 (2%)	8 (4%)
Other	9 (4%)	10 (4%)	2 (1%)	8 (4%)

Source: Reviewer analysis

Table 5 presents the baseline ISGA score and lesion counts by treatment arm. The baseline ISGA scores were fairly balanced between the two arms in both studies. The proportions of ISGA scores of 3 were slightly larger in the YAZ arms than the placebo arms, whereas the proportions of ISGA scores of 4 were marginally larger in the placebo arms compared to the YAZ arms in both studies. The mean baseline inflammatory and non-inflammatory lesion counts of the YAZ and placebo arms are very close in both studies. However, in Study 306820, the

maximum number of lesion counts (both inflammatory and non-inflammatory) are greater in the YAZ arm. There were 2 subjects that had more than 200 non-inflammatory lesions and 3 subjects with more than 100 inflammatory lesions at baseline in the YAZ arm. All subjects had at least 20 inflammatory and non-inflammatory lesions at baseline.

Table 5: Baseline Severity by Treatment Arm

	Study 306820		Study 306996	
	YAZ	Placebo	YAZ	Placebo
	N=228	N=230	N=218	N=213
ISGA				
3	134 (59%)	128 (56%)	134 (61%)	125 (59%)
4	81 (36%)	91 (40%)	74 (34%)	76 (36%)
5	14 (6%)	8 (3%)	14 (6%)	14 (7%)
Inflammatory lesion counts				
mean (std)	32.6 (16.1)	32.8 (14.6)	31.7 (12.4)	31.8 (13.7)
median	27	26.5	28	27
min,max	(20,152)	(20,100)	(20,98)	(20,104)
Non-inflammatory lesion counts				
mean (std)	47.3 (31.4)	47.0 (30.8)	43.9 (22.9)	43.8 (25.9)
median	36	36	37	36
min,max	(20,256)	(20,194)	(20,143)	(20,133)

Source: Reviewer analysis

3.1.4 Primary Efficacy Endpoints

3.1.4.1 ITT Analyses

The protocol defines the primary efficacy endpoints at Day 15±3 of Treatment Cycle 6 as the following:

- Percentage of subjects classified as 0 or 1 on the 6-grade ISGA scale;
- Percent change from Baseline in inflammatory lesion count;
- Percent change from Baseline in non-inflammatory lesion count;
- Percent change from Baseline in total lesion count.

Note that, as described in Section 3.1.1., the sponsor added visit windows in the SAP, dated 11/14/03 and that this addition was not submitted to the Agency. The primary analysis, which is based on the modified population, used these visit windows. Sensitivity analyses based on

the amended ITT population used all observations regardless of the date of measurement. The sponsor's analysis, based on the amended ITT population with visit windows, is also provided.

Table 6 presents the success rates based on the ISGA scores in the modified population. At Treatment Cycle 6, the primary endpoint, 15% and 21% of YAZ subjects reached success status, while 4% and 9% of placebo subjects were successes in Studies 306820 and 306996, respectively. The differences in the success rates of YAZ and placebo, in both studies, were highly statistically significant with p-values less than 0.0003 in both studies.

Table 6: ISGA Results (Modified Population)

	Study 306820		Study 306996	
	YAZ N=228	Placebo N=230	YAZ N=218	Placebo N=213
Number (Proportion) of Successes	35 (15.4%)	10 (4.3%)	46 (21.1%)	19 (8.9%)
p-value*		0.0001		0.0003

* p-values are calculated using a logistic model with treatment and pooled centers as factors

Source: Reviewer analysis based on modified population which excludes all subjects with Baseline ISGA score of 2 and includes all subjects with Baseline inflammatory/non-inflammatory lesion counts greater or equal to 20.

Table 7 presents the results of the other co-primary endpoints: percent change from Baseline to Treatment Cycle 6 in (i)inflammatory, (ii) non-inflammatory, and (iii) total lesion counts. In Study 306820, the difference in percent change from Baseline to Treatment Cycle 6 in all three lesion counts were highly statistically significant with p-values less than 0.0001. The ANCOVA model used to test the difference in the percent change of inflammatory and total lesion counts included Baseline lesion count as the covariate and treatment, center, and treatment by center interaction term as factors. The ANCOVA model used to test the percent change in non-inflammatory lesion counts included the same covariate and factors as the models used for inflammatory and total lesion counts, with the exception of the treatment by center interaction factor. This term was excluded from the model because it was tested to be statistically nonsignificant at the $\alpha = 0.10$ level. Treatment by center interaction in Study 306820 will be further discussed in Section 3.1.7. Study 306996 results replicated the findings from Study 306820. The differences in percent change in all three lesions counts were highly statistically significant with p-values less than 0.0001. The ANCOVA models used in Study 306996 included the same covariate and factors excluding the treatment by center interaction term.

The results from the sponsor's analyses are presented to investigate the impact of including subjects who had a baseline ISGA score of 2 on the 5-grade scale. Tables 8 and 9 are the sponsor's analyses of the co-primary endpoints.

Table 7: Mean (SD) Baseline and Percent Change in Lesion Counts (Modified Population)

	Study 306820		Study 306996	
	YAZ N=228	Placebo N=230	YAZ N=218	Placebo N=213
Inflammatory				
Baseline count	32.61 (16.1)	32.84 (14.6)	31.73 (12.4)	31.81 (13.7)
% reduction	47.60% (35.4%)	32.34% (37.3%)	50.60% (38.3%)	34.46 (49.7%)
p-value†		<0.0001		<0.0001
Non-inflammatory				
Baseline count	47.27 (31.4)	46.99 (30.6)	43.85 (22.9)	43.85 (25.9)
% reduction	38.08% (39.0%)	18.22% (47.9%)	42.35% (38.5%)	26.02% (48.2%)
p-value†		<0.0001		<0.0001
Total				
Baseline count	79.88 (42.3)	79.83 (37.2)	75.60 (30.7)	75.67 (33.9)
% reduction	42.33% (32.7%)	25.29% (36.4%)	46.13% (33.7%)	30.64% (41.9%)
p-value†		<0.0001		<0.0001

† p-values are calculated using ANCOVA model: Baseline lesion count as covariate, and treatment, center, and treatment by center interaction (if statistically significant) as factors.

Source: Reviewer analysis based on modified population which excludes all subjects with Baseline ISGA score of 2 and includes all subjects with Baseline inflammatory/non-inflammatory lesion counts greater or equal to 20.

Table 8: ISGA (Sponsor's Results)

	Study 306820		Study 306996	
	YAZ N=229	Placebo N=227	YAZ N=222	Placebo N=215
Number (Proportion) of Successes	37 (16.2%)	10 (4.4%)	47 (21.2%)	20 (9.3%)
p-value*		<0.0001†		0.0004‡

† P-value computed from logistic regression model with terms for treatment and pooled center.

‡ P-value computed from Cochran Mantel-Haenszel statistic stratified by pooled center, since the logistic regression model did not converge.

Source: Clinical Study Report No. A25083, p. 71 and Clinical Study Report No. A25152, p. 71

Compared to the modified population, the sponsor's analysis had two more successes in YAZ subjects in Study 306820, and one more successes on each arm in Study 306996. The difference in the results of percent change in lesion counts was very minor. All co-primary endpoints were highly statistically significant in the sponsor's results with p-values less than 0.0001. This analysis suggests that the statistically significant results were not driven by the subjects who were included with ISGA of 2.

Table 9: Mean (SD) Baseline and Percent Change in Lesion Counts (Sponsor's results)

	Study 306820		Study 306996	
	YAZ N=229	Placebo N=227	YAZ N=222	Placebo N=215
Inflammatory				
Baseline count	32.6 (16.1)	33.1 (14.6)	31.7 (12.3)	31.8 (13.6)
% reduction	47.8% (35.3%)	32.7% (36.4%)	50.9% (38.1%)	34.7 (49.6%)
p-value†		<0.0001		<0.0001
Non-inflammatory				
Baseline count	47.1 (31.4)	47.0 (30.8)	43.9 (22.8)	43.8 (25.8)
% reduction	38.4% (39.1%)	18.2% (48.7%)	42.8% (38.3%)	26.3% (48.1%)
p-value†		<0.0001		<0.0001
Total				
Baseline count	79.6 (42.3)	80.0 (37.4)	75.6 (30.5)	75.6 (33.7)
% reduction	42.6% (32.6%)	25.4% (36.7%)	46.5% (33.5%)	30.9% (41.8%)
p-value†		<0.0001		<0.0001

† p-values are calculated using ANCOVA model: Baseline lesion count as covariate, and treatment, center as factors.

Source: Clinical Study Reports No. A25083 and No. A25152, pages 60, 61, 65, 68 and Tables 22, 30, 38.

The modified population and sponsor's analyses considered observations assessed outside the visit window as missing. Since this window was not defined in the SAP reviewed by the Agency, this reviewer analyzed the co-primary endpoints based on the amended ITT population, which used all observed measurements, as defined in the protocol. Tables 10 and 11 presents the results based on the amended ITT population.

Compared to the modified population, the sponsor's analysis had four more successes in YAZ subjects and two more in placebo in Study 306820, and two more successes on each arm in Study 306996. The difference in the results of percent change in lesion counts was minor, with the amended ITT population percent changes being marginally larger than those of the modified population. All co-primary endpoints were highly statistically significant with p-values less than

Table 10: ISGA Results (Amended ITT Population)

	Study 306820		Study 306996	
	YAZ N=229	Placebo N=227	YAZ N=222	Placebo N=215
Number (Proportion) of Successes	39 (17.0%)	12 (5.3%)	48 (21.6%)	21 (9.8%)
p-value*		<0.0001		<0.0001

* p-values are calculated using a logistic model with treatment and pooled centers as factors

Source: Reviewer analysis based on sponsor's population disregarding visit windows

Table 11: Mean (SD) Baseline and Percent change in lesion counts (Amended ITT Population)

	Study 306820		Study 306996	
	YAZ N=229	Placebo N=227	YAZ N=222	Placebo N=215
Inflammatory				
Baseline count	32.55 (16.1)	33.07 (14.6)	31.68 (12.3)	31.79 (13.6)
% reduction	48.55% (35.8%)	33.84% (37.4%)	50.95% (40.3%)	36.06 (48.4%)
p-value†		<0.0001		0.0004
Non-inflammatory				
Baseline count	47.09 (31.4)	46.96 (30.8)	43.82 (22.8)	43.83 (25.8)
% reduction	39.50% (38.6%)	20.43% (47.4%)	43.59% (38.8%)	27.37% (48.4%)
p-value†		<0.0001		<0.0001
Total				
Baseline count	79.64 (42.3)	80.03 (37.4)	75.51 (30.5)	75.62 (33.7)
% reduction	44.00% (32.0%)	27.02% (37.0%)	46.86% (34.6%)	31.74% (41.8%)
p-value†		<0.0001		<0.0001‡

† p-values are calculated using ANCOVA model: Baseline lesion count as covariate, and treatment, center, and treatment by center interaction (if statistically significant) as factors.

Source: Reviewer analysis based on sponsor's population disregarding visit windows

0.0004. This analysis suggests that the statistically significant results were not driven by the use of visit windows. The results from the three analyses above confirmed that the difference in ISGA success rates and percent changes in lesion counts of the two arms were highly significant in both studies.

3.1.4.2 Sensitivity Analysis of the Primary Efficacy Endpoint

Per protocol, last observation carried forward (LOCF) was used to impute missing data in the analyses of the previous section. The details of the numbers and proportions of missing observations in each treatment arm over time is provided in the Appendix A.1.

The submission nor protocol had included a sensitivity analysis method to ensure that the results were not driven by the imputation method. This reviewer used two imputation methods as sensitivity analyses based on the modified population. The first method imputed all missing ISGA observations as successes and used each arm's mean percent change of the subjects who were successes based on their ISGA score to impute the missing lesion count data (Table 12). For ISGA, the results in Tables 12 and 13 are the number (proportion) subjects who reached success status, and for lesion counts the results are the mean percent reduction in lesion counts.

Table 12: Sensitivity Analysis (Impute all missing as success)

	Study 306820			Study 306996		
	YAZ N=228	Placebo N=230	p-value	YAZ N=218	Placebo N=213	p-value
Imputed subjects	54 (24%)	60 (26%)		51 (23%)	61 (28%)	
ISGA§	86 (38%)	70 (30%)	0.0934‡	97 (44%)	85 (40%)	0.3023‡
Inflammatory lesions§§	63% (29%)	50% (38%)	<0.0001†	68% (30%)	56% (48%)	0.0126†
Non-inflammatory lesions§§	52% (32%)	34% (44%)	<0.0001†	58% (35%)	44% (47%)	0.0001†

§ Number (proportion) of successes

§§ Mean (SD) percentage change from baseline

‡ p-value was calculated using logistic regression with treatment and center as factors.

† p-value was calculated using ANCOVA model: baseline lesion count as covariate and treatment, center, and treatment by center interaction (if statistically significant) as factors.

Source: Reviewer analysis

When all missing observations were imputed as successes, the difference in the ISGA scores were not statistically significant with p-values of 0.0934 and 0.3023, for Studies 306820 and 306996, respectively. This approach is very conservative. There were 48 and 50 subjects in YAZ arm who did not reach success at Cycle 3 and were imputed as success at Cycle 6, compared to 60 and 69 subjects in the placebo arm. The placebo arm had more missing data and more subjects who discontinued the study due to lack of efficacy. Also, considering that YAZ had a higher success rate than placebo, it is more likely that the proportion of success among subjects who are missing at Cycle 6 would be higher in YAZ than placebo. The differences in percent change in lesion counts were statistically significant using this imputation method for all three types of lesions.

Table 13: Sensitivity Analysis (Impute all missing as failures)

	Study 306820			Study 306996		
	YAZ N=228	Placebo N=230	p-value	YAZ N=218	Placebo N=213	p-value
Imputed subjects	54 (24%)	60 (26%)		51 (23%)	61 (28%)	
ISGA§	32 (14%)	10 (4%)	0.0005‡	45 (21%)	18 (8%)	0.0001‡
Inflammatory lesions§§	55% (27%)	38% (33%)	<0.0001†	59% (28%)	38% (45%)	<0.0001†
Non-inflammatory lesions§§	44% (31%)	24% (41%)	<0.0001†	48% (33%)	29% (43%)	<0.0001†

§ Number (proportion) of successes

§§ Mean (SD) percentage change from baseline

‡ p-value was calculated using CMH test stratified by pooled sites

† p-value was calculated using ANCOVA model: baseline lesion count as covariate and treatment, center as factors.

Source: Reviewer analysis

The second method imputes all missing ISGA scores as failures and correspondingly imputes missing lesion count data as the mean percent change of the subjects who were failures according to their ISGA score for each treatment (Table 13). The results when imputing all missing observations as failures are similar to that of the LOCF approach. Compared to the modified population results, YAZ arm has three less successes in Study 306820 when using this imputation method. Nonetheless, all efficacy endpoints are highly statistically significant when missing observations are imputed as failures, with p-values less than 0.0005.

In addition to the sensitivity analyses regarding the imputation methods, this reviewer compared the proportion of successes, based on ISGA, by treatment group in subjects who did not meet the modified inclusion criteria (Table 14).

Table 14: ISGA Results (Excluded Subjects)

	Study 306820		Study 306996	
	YAZ N=38	Placebo N=38	YAZ N=52	Placebo N=55
Number (Proportion) of Successes	7 (18%)	2 (5%)	16 (31%)	7 (13%)

Source: Reviewer analysis

Subjects who were excluded from the modified population had higher success rates than those who were included in each treatment arm for both studies (See Table 6). The number

of the excluded subjects were not sufficient to draw statistical inference. Nonetheless, these results suggest that the difference between the modified population and excluded subjects is not substantial.

3.1.4.3 Per Protocol Analysis

The per protocol (PP) population included subjects who met the amended inclusion criteria, had 80% or higher overall study drug compliance, had no major protocol violations, and completed a minimum of 5 treatment cycles. A total of 310 subjects were excluded from the per protocol population, 167 subjects (36%) from the YAZ arm and 148 subjects (34%) from the placebo arm. Table 15 presents the results of the primary endpoint analyses at Treatment Cycle 6 on the per protocol population.

Table 15: Per Protocol Population Analysis-Number (Proportion) of Success on ISGA and Mean Baseline and % Reduction (SD) in Lesion Counts

	Study 306820		Study 306996	
	YAZ N=148	Placebo N=143	YAZ N=145	Placebo N=138
ISGA				
Number of Successes	31 (21%)	8 (6%)	40 (28%)	16 (12%)
p-value‡		0.0003		0.0010
Inflammatory lesions				
% reduction	60.2% (28.0%)	39.8% (38.0%)	60.0% (33.1%)	41.2% (51.5%)
p-value†		<0.0001		0.0005
Non-inflammatory lesions				
% reduction	48.5% (34.2%)	23.4% (50.9%)	48.8% (37.9%)	30.3% (52.1%)
p-value†		<0.0001		0.0010
Total lesions				
% reduction	53.9% (26.1%)	31.7% (38.5%)	54.0% (30.7%)	35.7% (44.6%)
p-value†		<0.0001		0.0001

‡ p-values are calculated using logistic regression with treatment and center as factors.

† p-values are calculated using ANCOVA model: Baseline lesion count as covariate, and treatment, center as factors.

Source: Reviewer analysis

The proportion of successes based on the ISGA score were higher in the PP population than the modified population for both arms in both studies. The differences in the proportion of successes of the two arms were statistically significant with p-values of less than 0.0003 and 0.0010 in studies 306820 and 306996, respectively. The differences in percent change in the

lesion count from baseline were also statistically significant with p-values less than 0.0010. The modified and PP populations primary endpoint analyses results were similar, which further supports the superiority of YAZ over placebo.

3.1.5 Secondary Efficacy Endpoints

The secondary endpoints defined in the protocol were

- Change from Baseline in individual lesion counts (open comedones, closed comedones, papules, pustules, nodules);
- Percentage of subjects classified as "improved" according to the Investigator's Overall Improvement Rating;
- Percentage of subjects classifying themselves as "improved" on the Subject's Overall Self-Assessment Rating.

As indicated in Section 3.1.1 the Agency conveyed comments to the sponsor that the Investigator's Overall Improvement Rating and the Subject's Overall Self-Assessment Rating have little regulatory utility. Also, that multiplicity adjustment would be needed for labeling claim. The other secondary endpoints, change from Baseline in individual lesion counts are not relevant to labeling claims. Although not defined as secondary endpoints by the sponsor, this reviewer analyzed the absolute change in lesion counts for inflammatory, non-inflammatory, and total lesion counts. Table 16 presents the results.

The differences in absolute changes of lesion counts were all highly statistically significant in all lesion types with p-values less than 0.0001. Tables 23 and 24 in the Appendix A.2 present the sponsor's analysis results of the secondary endpoints.

Table 16: Mean (SD) Baseline and Absolute Change in Lesion Counts (Modified Population)

	Study 306820		Study 306996	
	YAZ N=228	Placebo N=230	YAZ N=218	Placebo N=213
Inflammatory				
Baseline count	32.61 (16.1)	32.84 (14.6)	31.73 (12.4)	31.81 (13.7)
Absolute change	15.33 (13.8)	10.93 (13.7)	15.80 (13.8)	10.77 (15.6)
p-value†		<0.0001		<0.0001
Non-inflammatory				
Baseline count	47.27 (31.4)	46.99 (30.6)	43.85 (22.9)	43.85 (25.9)
Absolute change	18.02 (23.0)	9.83 (23.3)	17.30 (17.9)	10.91 (19.4)
p-value†		<0.0001		<0.0001
Total				
Baseline count	79.88 (42.3)	79.83 (37.2)	75.60 (30.7)	75.67 (33.9)
Absolute change	33.35 (31.4)	20.77 (31.1)	33.10 (26.6)	21.68 (29.4)
p-value†		<0.0001		<0.0001

† p-values are calculated using ANCOVA model: Baseline lesion count as covariate, and treatment, center, and treatment by center interaction (if statistically significant) as factors.

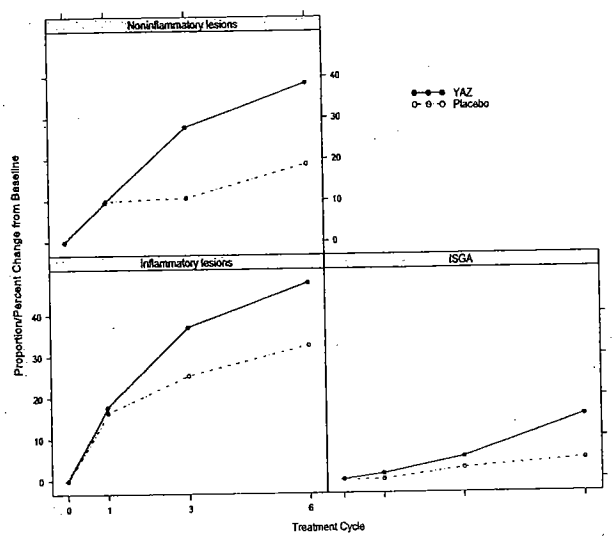
Source: Reviewer analysis based on the modified population which excludes all subjects with baseline ISGA score of 2 and includes all subjects with baseline inflammatory/non-inflammatory lesion counts greater or equal to 20.

3.1.6 Efficacy Results over Time

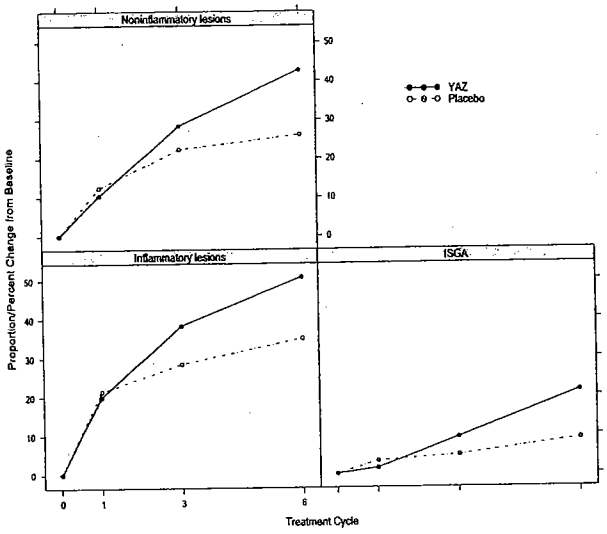
Subjects were treated and followed for 24 weeks (168 days). The subject's ISGA score and lesion counts were evaluated at Baseline, Day 15±3 of Cycles 1, 3, and 6. Figure 1 presents the success rates based on the ISGA scores, percent change in inflammatory and non-inflammatory lesion counts from baseline over time. Note that analyses of the total lesion counts will be excluded from this review hereafter, since the total lesion count is the sum of inflammatory and non-inflammatory lesion counts.

The success rates based on ISGA scores are similar at the end of Cycle 1. However, the treatment effect increases as the treatment duration increases. This trend is also true for both inflammatory and non-inflammatory lesion counts.

Figure 1: Efficacy Over Time



(a) Study 306820



(b) Study 306996

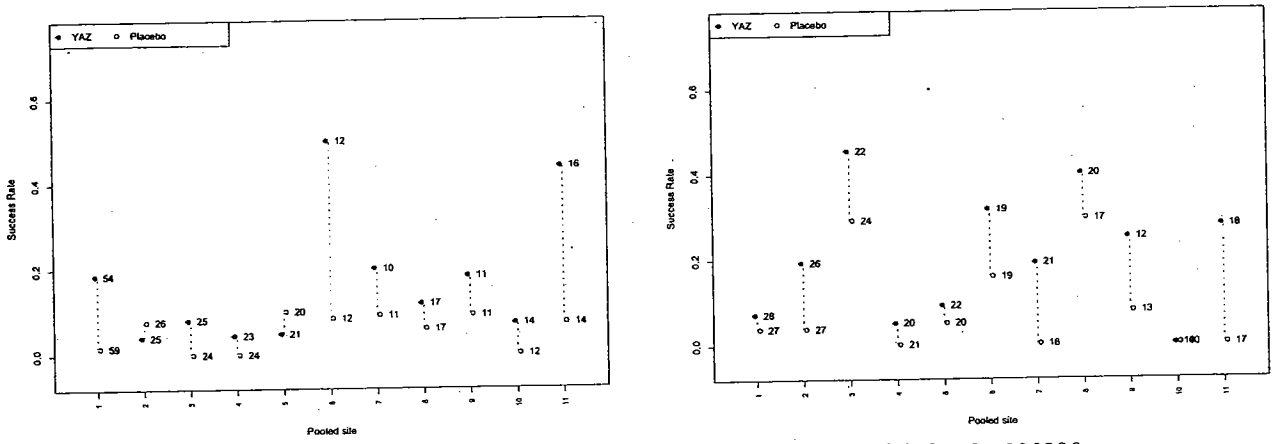
3.1.7 Efficacy Results by Center

The modified population of this submission enrolled subjects from 24 and 22 investigative sites, all from the United States, for Studies 306820 and 306996, respectively. The maximum numbers of enrollment by one site were 113 and 55 subjects in these studies. Study center [redacted] enrolled approximately 24% of the total number of subjects in Study 306820. To evaluate the impact of this extremely large center on the success rate based on the ISGA score, this reviewer conducted the analysis after excluding this center (see Appendix A.3, Table 25). Even after excluding the 113 subjects from the analysis, the difference in the two arms was highly statistically significant with a p-value of 0.0049. Sites that did not meet the target number of subjects, 10 subjects per arm per site, were pooled from the largest to the smallest until the pooled center had the target number of subjects. Study sites were pooled into 11 pooled centers in both studies.

Figure 2 presents the success rates based on the ISGA scores and number of subjects enrolled in each pooled site by treatment. The success rates of both arms in both studies appeared to be relatively consistent across the pooled sites, and therefore the results do not seem to be driven by extreme sites. The Breslow-Day test results also supported this conclusion with p-values of 0.3976 and 0.7536 for studies 306820 and 306996, respectively.

Figure 3 presents the percent change in inflammatory and non-inflammatory lesion counts from Baseline at Treatment Cycle 6 by site. Figures 3 (a) and (c) present the results from Study 306820. Site 6 (200188 [redacted] appeared to have a substantially larger treatment effect compared to other sites in both lesion types. The treatment by center interaction

Figure 2: ISGA by Site



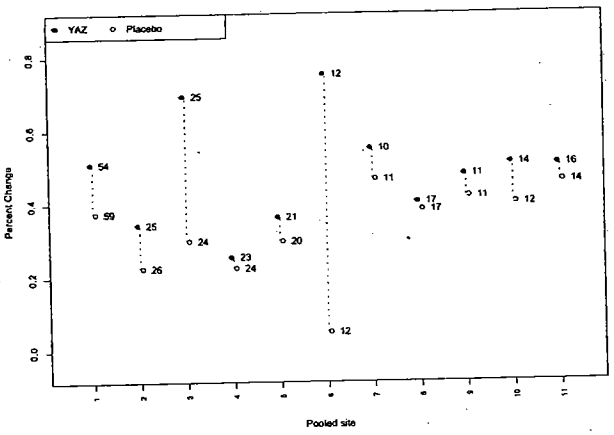
(a) Study 306820

(b) Study 306996

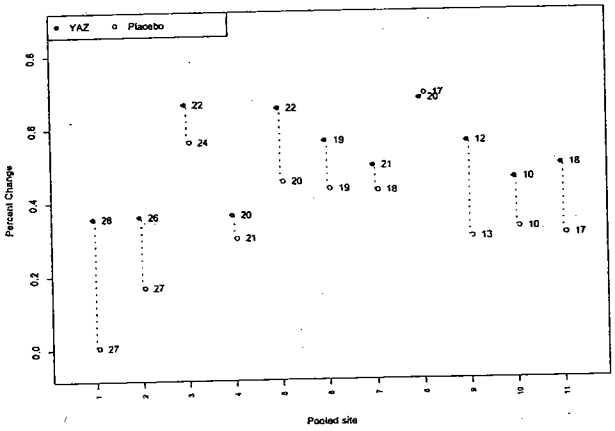
term in the ANCOVA model that was used to compare the two arms was statistically significant (p-value To investigate whether the efficacy results were driven by this extreme center, a sensitivity analysis that excluded Site 6 was done. (See Appendix A.3, Table 27.) The treatment by center interaction term was no longer statistically significant after Site 6 was excluded from the analysis (p-value The difference in the percent change from baseline were statistically significant for both lesions with p-values of 0.0003 and <0.0001 in inflammatory and non-inflammatory lesions, respectively. (The difference in the success rate based on ISGA score remained statistically significant after excluding this site with a p-value of 0.0010, Appendix A.3, Table 26.) Based on the results of this sensitivity analysis, we conclude that the efficacy results were highly statistically significant even without the extreme center (Site 6) and that the efficacy results were not driven by this site.

Figures 3 (b) and (d) present the results from Study 306996. The treatment effect on both lesion types appeared to be relatively consistent across the pooled sites and therefore the results do not seem to be driven by extreme sites.

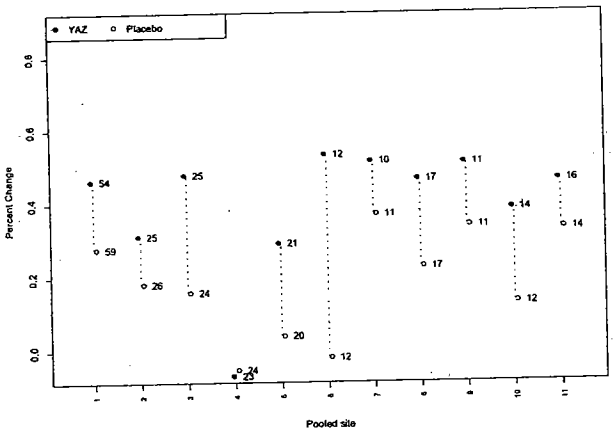
Figure 3: Percent Change in Lesion Counts from Baseline by Sites



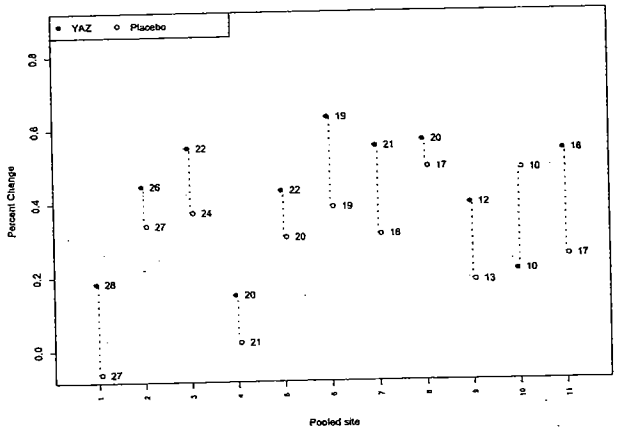
(a) Inflammatory: Study 306820



(b) Inflammatory: Study 306996



(c) Non-inflammatory: Study 306820



(d) Non-inflammatory: Study 306996

3.2 Evaluation of Safety

The safety evaluation includes all patients that were randomized and dispensed study medication, regardless of whether the inclusion criteria was met. There were a total of 534 subjects evaluated for safety in Study 306820 (YAZ: 266 , Placebo: 268) and 538 subjects in Study 306996 (YAZ: 270, Placebo: 268). This section includes the extent of drug exposure and adverse events.

3.2.1 Extent of Exposure

The duration of treatment was defined as (date of last treatment)-(date of first treatment)+1. Most subjects in both arms used the treatment for 168 days (6 cycles). In both studies and both arms, the median duration of treatment was 168 days. In Study 306820, the mean treatment duration was 141.7 days (range 0−196 days) in the YAZ arm and 139.0 days (range 0−231 days) in the placebo arm. The mean treatment duration in Study 306996 was slightly longer than that

of Study 306820, with means of 146.4 days (range 0−207 days) and 144.1 days (range 0−218 days) in the YAZ and placebo arms, respectively.

3.2.2 Adverse Events

A total of 299 (56.0%) and 272 (50.6%) subjects in Studies 306820 and 306996 reported at least 1 treatment-emergent adverse event (AE). The proportion of subjects who experienced such AE was higher in the YAZ arm than the placebo: 63.9% and 48.1% of YAZ and placebo subjects, respectively, in Study 306820, and 55.2% and 45.9% of YAZ and placebo subjects in Study 306996. Table 17 present AE rates for events occurring at least 3% of subjects per treatment arm.

Table 17: Adverse Events Occurring in at Least 3% of Subjects per Treatment Arm

	Study 306820		Study 306996	
	YAZ	Placebo	YAZ	Placebo
Preferred Term	N=266	N=268	N=270	N=268
Subjects with at least 1 AE	170 (64%)	129 (48%)	149 (55%)	123 (46%)
Upper respiratory infection	35 (13%)	25 (9%)	25 (9%)	31 (12%)
Metrorrhagia	28 (11%)	3 (1%)	25 (9%)	6 (2%)
Headache	23 (9%)	10 (4%)	14 (5%)	15 (5%)
Nausea	17 (6%)	7 (3%)	9 (3%)	7 (3%)
Sinusitis	15 (6%)	9 (3%)	5 (2%)	5 (2%)
Pap smear suspicious	15 (6%)	12 (5%)	15 (6%)	8 (3%)
Vaginal moniliasis	15 (6%)	12 (5%)	5 (2%)	4 (2%)
Flu syndrome	8 (3%)	3 (1%)	10 (4%)	12 (5%)
Tooth disorder	8 (3%)	3 (1%)	1 (<1%)	3 (1%)
Emotional lability	8 (3%)	6 (2%)	4 (2%)	2 (1%)
Menorrhagia	8 (3%)	2 (1%)	8 (3%)	3 (1%)
Abdominal pain	7 (3%)	1 (<1%)	4 (2%)	6 (2%)
Depression	7 (3%)	5 (2%)	5 (2%)	0 (0%)
Breast pain	7 (3%)	1 (<1%)	1 (<1%)	2 (1%)
Dysmenorrhea	7 (3%)	3 (1%)	1 (<1%)	1 (<1%)
Pharyngitis	4 (2%)	10 (4%)	4 (2%)	4 (2%)
Menstrual disorder	1 (<1%)	1 (<1%)	7 (3%)	2 (1%)

Source: Clinical Study Report No. A25083, p. 82 and Clinical Study Report No. A25152, p. 83

The most common adverse event was upper respiratory tract infection, which occurred in approximately 9% of the total subjects. This AE rate was higher in the YAZ arm than the placebo in Study 306820, however opposite results were shown in Study 306996. The next most

common AE was metrorrhagia (10% of YAZ and 2% of placebo subjects). Two subjects in Study 306820 reported serious adverse events (SAE). One subject on the YAZ group reported depression, which was assessed by the investigator as unlikely related to study drug. One subject on the placebo arm reported gastrointestinal disorder, which was assessed by the investigator as not related to study drug. In Study 306996, 7 SAEs were reported. One subject on YAZ had pneumonia, which was assessed as not related to the study drug. The other SAEs occurred in subjects on the placebo arm. The following reported SAEs were assessed as not related to the study drug by the investigator, alcohol abuse, pelvic pain, ectopic pregnancy, abdominal pain, accidental injury. One subject had epileptic seizure, which was assessed as unlikely related to study drug. No deaths were reported in either studies.

4 Findings in Special/Subgroup Populations

4.1 Race and Age

Table 18 presents the ISGA success rates by race and age. The success rate in YAZ arm is higher than that of placebo arm for all subgroups in Study 306820. This is also true for Study 306996 with the exception of Hispanics, where placebo subjects had a marginally higher success rate than YAZ subjects. Some subgroups did not have sufficient number of subjects to draw any reasonable inference. The success rates are relatively consistent across age groups in both studies. With the exception of one subgroup (Age 23 - 26) in Study 306996, the success rates of YAZ groups were higher than that of the placebo groups in all age categories in both studies. The subgroup analysis supports the claim that the success rates (ISGA) in YAZ subjects are higher than that of placebo across subgroups.

Table 18: ISGA (Number (%) of Successes) by Race and Age

			Study 306820		Study 306996	
			YAZ N=228	Placebo N=230	YAZ N=218	Placebo N=213
Race	Caucasian	Total	161	146	157	142
		Success (%)	31 (19%)	8 (5%)	36 (23%)	12 (8%)
	Black	Total	31	41	34	27
		Success (%)	1 (3%)	0 (0%)	5 (15%)	3 (11%)
	Hispanic	Total	23	26	20	28
		Success (%)	3 (13%)	2 (8%)	2 (10%)	3 (11%)
	Asian	Total	4	7	5	8
		Success (%)	0 (0%)	0 (0%)	2 (40%)	1 (13%)
	Other	Total	9	10	2	8
		Success (%)	0 (0%)	0 (0%)	1 (50%)	0 (0%)
Age	14 - 22	Total	102	91	97	98
		Success (%)	14 (14%)	3 (3%)	24 (25%)	3 (3%)
	23 - 26	Total	53	52	43	48
		Success (%)	8 (15%)	3 (6%)	4 (9%)	7 (15%)
	27 - 30	Total	29	37	26	25
		Success (%)	4 (14%)	3 (8%)	6 (23%)	3 (12%)
	31 - 34	Total	23	26	25	12
		Success (%)	4 (17%)	1 (4%)	6 (24%)	1 (8%)
	35 - 45	Total	21	24	27	30
		Success (%)	5 (23%)	0 (0%)	6 (22%)	5 (17%)

Source: Reviewer analysis

Table 19 presents the mean (sd) percent change from baseline in inflammatory and non-inflammatory lesion counts by race and age. The mean percent changes in inflammatory and non-inflammatory lesion count were relatively consistent over the races within each study. The mean percent changes in lesion counts were consistent over age categories in both studies for both lesion types.

Table 19: Mean (SD) Percent Change in Lesion Counts by Race and Age

		Study 306820				Study 306996			
		Inflammatory		Noninflammatory		Inflammatory		Noninflammatory	
		YAZ	Placebo	YAZ	Placebo	YAZ	Placebo	YAZ	Placebo
		N=228	N=230	N=228	N=230	N=218	N=213	N=218	N=213
Race	Caucasian	n=161 51% (35%)	n=146 32% (39%)	n=161 42% (37%)	n=146 26% (44%)	n=157 52% (39%)	n=142 33% (50%)	n=157 47% (37%)	n=142 30% (46%)
	Black	n=31 40% (31%)	n=41 36% (28%)	n=31 21% (30%)	n=41 2% (47%)	n=34 50% (32%)	n=27 38% (43%)	n=34 30% (39%)	n=27 18% (33%)
	Hispanic	n=23 47% (33%)	n=26 29% (45%)	n=23 33% (51%)	n=26 14% (57%)	n=20 41% (41%)	n=28 29% (60%)	n=20 27% (48%)	n=28 10% (69%)
	Asian	n=4 -6% (12%)	n=7 27% (29%)	n=4 17.0% (34%)	n=7 -9.0% (40%)	n=5 56% (46%)	n=8 61% (28%)	n=5 53% (36%)	n=8 32% (57%)
	Other	n=9 33% (47%)	n=10 41% (32%)	n=9 40% (31%)	n=10 -1% (62%)	n=2 62% (4%)	n=8 42% (32%)	n=2 54% (37%)	n=8 30% (25%)
Age	14 - 22	n=102 52% (35%)	n=91 25% (39%)	n=102 41% (31%)	n=91 12% (50%)	n=97 53% (36%)	n=98 27% (48%)	n=97 38% (40%)	n=98 25% (49%)
	23 - 26	n=53 49% (33%)	n=52 31% (40%)	n=53 36% (48%)	n=52 20% (48%)	n=43 36% (44%)	n=48 32% (66%)	n=43 36% (37%)	n=48 23% (55%)
	27 - 30	n=29 39% (33%)	n=37 36% (31%)	n=29 30% (48%)	n=37 10% (54%)	n=26 61% (32%)	n=25 46% (32%)	n=26 50% (41%)	n=25 26% (34%)
	31 - 34	n=23 34% (46%)	n=26 41% (32%)	n=23 31% (41%)	n=26 30% (38%)	n=25 46% (45%)	n=12 37% (34%)	n=25 48% (38%)	n=12 29% (43%)
	35 - 45	n=21 51% (33%)	n=24 49% (32%)	n=21 51% (32%)	n=24 35% (35%)	n=27 60% (31%)	n=30 53% (36%)	n=27 55% (29%)	n=30 32% (48%)

Source: Reviewer analysis

4.2 Other Special/Subgroup Populations

The proportion of success rates based on the ISGA score and the mean percent change in lesion counts were explored by Baseline disease severity (Baseline ISGA score). Tables 20 and 21 present the success rates and mean percent change in lesion counts from baseline across baseline ISGA scores. Subjects who entered the study with moderate severity (ISGA score of 3) had higher success rates than subjects with more severe disease at baseline (ISGA score of 4 or 5) in both arms and studies. However, the number of subjects in this severity group is too small to draw any reasonable inference.

Table 20: ISGA Number (%) of Successes by Baseline Disease Severity

			Study 306820		Study 306996	
			YAZ N=228	Placebo N=230	YAZ N=218	Placebo N=213
Baseline ISGA	3	Total	134	128	134	125
		Success (%)	31 (23%)	8 (6%)	37 (28%)	17 (14%)
	4	Total	81	91	74	76
		Success (%)	7 (9%)	4 (4%)	11 (15%)	3 (4%)
	5	Total	14	8	14	14
		Success (%)	1 (7%)	0 (0%)	0 (0%)	1 (7%)

Source: Reviewer analysis

Table 21: Percent Change in Lesion Counts by Baseline Disease Severity

			Study 306820				Study 306996			
			Inflammatory		Noninflammatory		Inflammatory		Noninflammatory	
			YAZ N=228	Placebo N=230	YAZ N=228	Placebo N=230	YAZ N=218	Placebo N=213	YAZ N=218	Placebo N=213
Baseline ISGA	3	n	134	128	134	128	134	125	134	125
		Mean	48%	32%	37%	16%	51%	34%	45%	28%
		S.D.	(39%)	(38%)	(43%)	(45%)	(40%)	(56%)	(40%)	(55%)
	4	n	81	91	81	91	74	76	74	76
		Mean	50%	33%	42%	21%	52%	33%	41%	22%
		S.D.	(29%)	(37%)	(32%)	(52%)	(36%)	(41%)	(36%)	(39%)
	5	n	14	8	14	8	14	14	14	14
		Mean	27%	30%	25%	18%	43%	49%	25%	31%
		S.D.	(32%)	(29%)	(35%)	(37%)	(38%)	(33%)	(36%)	(28%)

Source: Reviewer analysis

5 Summary and Conclusions

5.1 Statistical Issues and Collective Evidence

The sponsor conducted two studies (Study 306820 and Study 306996) under the protocol that was agreed upon with the Agency in terms of study design and endpoints. Efficacy was evaluated on Day 15 of Treatment Cycle 6 for the proportion of success rate based on the ISGA score and the percent change in lesion (inflammatory, non-inflammatory, and total) counts from baseline. The difference in the success rates were statistically significant in both studies (p-values <0.0003). The difference in the mean percent changes in lesion counts were also statistically significant in all lesion types in both studies (p-values <0.0001). Within each study, the efficacy results were relatively consistent across subgroups and investigative sites. However, in Study 306820, site 6 (200188=[redacted]) showed larger treatment effects compared to that of other study sites. Sensitivity analyses which excluded this extreme center still demonstrated superiority of YAZ tablets to placebo. The protocol defined seven secondary endpoints, which had little regulatory utility or were not relevant to labeling claims.

5.2 Conclusions and Recommendations

YAZ tablets has been demonstrated to be statistically superior to placebo in two studies (Study 306820 and Study 309669) in the treatment of moderate acne. Efficacy was evaluated on (i) success rate based on the Investigator Static Global Assessment (ISGA) score and (ii) percent change from baseline for two out of three lesions counts (inflammatory, non-inflammatory, and total) at Day 15 of Treatment Cycle 6. Table 22 presents the summary of the co-primary endpoint results. All co-primary endpoints were statistically significant in both studies with p-values less than 0.0003.

The adverse event rates were higher in YAZ subjects than in the placebo subjects. The most common adverse event was upper respiratory tract infection, which was reported by approximately 9% of the subjects. The next common adverse event was metrorrhagia.

Table 22: Efficacy Results Summary - Number (Proportion) of success on ISGA and Mean Baseline and % Reduction (SD) in Lesion Counts

	Study 306820		Study 306996	
	YAZ N=228	Placebo N=230	YAZ N=218	Placebo N=213
ISGA				
Number of Successes	35 (15.4%)	10 (4.3%)	46 (21.1%)	19 (8.9%)
p-value*		0.0001		0.0003
Inflammatory				
Baseline count	32.61 (16.1)	32.84 (14.6)	31.73 (12.4)	31.81 (13.7)
% reduction	47.60% (35.4%)	32.34% (37.3%)	50.60% (38.3%)	34.46 (49.7%)
p-value†		<0.0001		<0.0001
Non-inflammatory				
Baseline count	47.27 (31.4)	46.99 (30.6)	43.85 (22.9)	43.85 (25.9)
% reduction	38.08% (39.0%)	18.22% (47.9%)	42.35% (38.5%)	26.02% (48.2%)
p-value†		<0.0001		<0.0001
Total				
Baseline count	79.88 (42.3)	79.83 (37.2)	75.60 (30.7)	75.67 (33.9)
% reduction	42.33% (32.7%)	25.29% (36.4%)	46.13% (33.7%)	30.64% (41.9%)
p-value†		<0.0001		<0.0001

* p-values are calculated using a logistic model with treatment and pooled centers as factors

† p-values are calculated using ANCOVA model: Baseline lesion count as covariate, and treatment, center, and treatment by center interaction (if statistically significant) as factors.

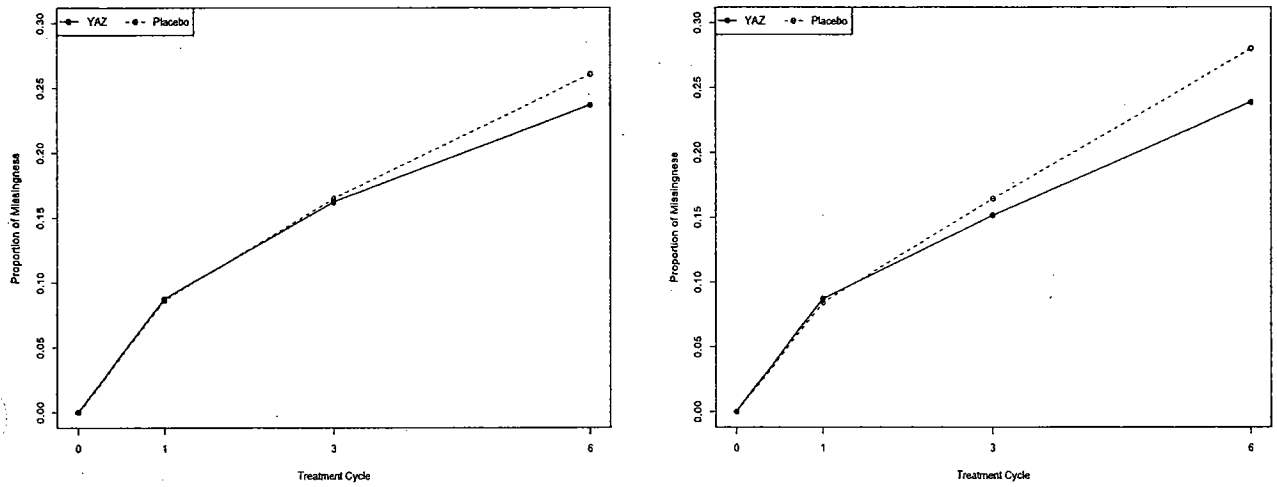
Source: Reviewer analysis

APPENDIX

A.1 Number and Proportion of Missing Observations

Figure 4 presents the number and proportion of missing observations in each treatment arm over time.

Figure 4: Proportion of Missing Observations Over Time



(a) Study 306820

(b) Study 306996

The number of missing observations at Treatment Cycle 6 were a total of 114 (25%) and 112 (28%) for Studies 306820 and 306996, respectively. The placebo arms had larger proportions of missing observations than the YAZ arms at Treatment Cycles 3 and 6, in both studies. At Treatment Cycle 6, 24% and 23% of the YAZ observations were missing in Studies 306820 and 306996, respectively, whereas in the placebo arms, 26% and 28% of the observations were missing in the two studies. The number of missing observations at Cycle 6 is relatively large because observations assessed after Day 168 were considered as missing.

A.2 Sponsor's Secondary Efficacy Endpoints Analyses

Table 23: Sponsor's Secondary Endpoint Analysis (Overall Rating)

	Study 306820		Study 306996	
	YAZ N=176	Placebo N=163	YAZ N=179	Placebo N=169
Investigator's Overall Improvement Rating				
Number (%) Who Improved	159 (89%)	115 (68%)	152 (86%)	104 (64%)
p-value†		<0.0001		<0.0001
Subject's Overall Self-Assessment Rating				
Number (%) Who Improved	159 (89%)	125 (74%)	157 (89%)	99 (61%)
p-value†		0.0005		<0.0001

† p-values were calculated from logistic regression model with terms for treatment and pooled centers

Source: Clinical Study Report No. A25083, p. 74 and Clinical Study Report No. A25152, p. 74

Table 24: Mean Change from Baseline at Endpoint

	Study 306820			Study 306996		
Type of lesion	YAZ N=229	Placebo N=227	p-value	YAZ N=222	Placebo N=215	p-value
Papules	10.1	7.3	0.0002†	10.2	6.8	0.0018†
Pustules	5.1	3.9	0.0169†	5.4	4.2	0.0286‡
Nodules	0.1	-0.1	0.0422‡	0.3	-0.1	0.0042‡
Open Comedones	7.7	4.9	0.0007††	8.2	4.9	0.0267‡
Closed Comedones	10.4	4.9	0.0002†	9.4	6.1	0.0005†

† p-value from ANCOVA with terms treatment, pooled center, and Baseline covariate

‡ p-value from ANOVA with terms treatment and pooled center

†† p-value from rank ANOVA with terms treatment and pooled center, as normality was rejected at 0.05 level.

Source: Clinical Study Report No. A25083, p. 74-78 and Clinical Study Report No. A25152, p. 74-78

A.3 Sensitivity Analysis

A.3.1 ISGA Results After Excluding Site 1

Table 25: ISGA Results (Site 1 Excluded)

	Study 306820		
	YAZ N=174	Placebo N=171	p-value*
Number (Proportion) of Successes	25 (14.4%)	8 (5.5%)	0.0049

* p-values are calculated using a logistic model with treatment and pooled centers as factors

Source: Reviewer analysis

A.3.2 Efficacy Results After Excluding Site 6

Table 26: ISGA Results (Site 6 Excluded)

	Study 306820		
	YAZ N=216	Placebo N=218	p-value*
Number (Proportion) of Successes	29 (13.4%)	9 (4.1%)	0.0010

* p-values are calculated using a logistic model with treatment and pooled centers as factors

Source: Reviewer analysis

Table 27: Mean (SD) Baseline and Percent Change in Lesion Counts (Site 6 Excluded)

	YAZ N=218	Placebo N=216	p-value*
Inflammatory lesion			
Baseline count	33.00 (16.3)	33.07 (14.7)	
% reduction	46.07% (35.3%)	33.84% (36.1%)	0.0003
Non-inflammatory lesion			
Baseline count	46.00 (30.7)	45.39 (29.4)	
% reduction	37.25% (39.7%)	19.33% (47.1%)	< 0.0001

* p-values are calculated using ANCOVA model: Baseline lesion count as covariate, and treatment and center as factors.

Source: Reviewer analysis

Signatures/Distribution List

Primary Statistical Reviewer: Clara Kim, Ph.D.
Date: December 1, 2006

Statistical Team Leader: Mohamed Alosh, Ph.D.

cc:
Archival NDA
DDDP/Walker
DDDP/Kukich
DDDP/Luke
DDDP/Kettl
DDDP/Wright
OBIO/O'Neill
OBIO/Patrician
DBIII/Wilson
DBIII/Alosh
DBIII/Kim
December 1, 2006

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

/s/

Clara Kim
12/1/2006 04:34:23 PM
BIOMETRICS

Mohamed Alosh
12/1/2006 04:40:53 PM
BIOMETRICS
Concur with review