CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

22-320

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:	22-320/SN000
Drug Name:	EPIDUO (benzoyl peroxide and adapalene)
Indication(s):	Acne vulgaris
Applicant:	Galderma
Dates:	Submitted: 02/08/2008 PDUFA: 12/08/2008
Review Priority:	Standard
Biometrics Division:	Division of Biometrics III
Statistics Reviewer:	Mat Soukup, Ph.D.
Concurring Reviewer:	Mohamed Alosh, Ph.D.
Medical Division:	Division of Dermatology and Dental Products
Clinical Team:	Reviewer: Jane Liedtka, M.D. (DDDP) Lead: Jill Lindstrom, M.D. (DDDP)
Project Manager:	Dawn Williams (DDDP)
Keywords:	combination product, superiority

Contents

1 Executive Summary

1.1 Conclusions and Recommendations

EPIDUO is a fixed combination product comprising adapalene 0.1% (w/w) and benzoyl peroxide 2.5% (w/w) in a topical gel for the treatment of acne vulgaris. The clinical program consisted of one Phase 2 study (Study 18094), one Phase 3 study (Study 18087), and one long-term open label safety study (Study 18089). Study 18094 and Study 18087 were used to assess the efficacy of EPIDUO as compared to each monad and its vehicle. The primary efficacy endpoints were:

- Change from baseline in inflammatory and non-inflammatory lesion counts.
- Percent of patients with an IGA success defined as 'clear' or 'almost clear'.

In both studies, EPIDUO was statistically superior to each monad and vehicle for the percent of IGA successes and the change in non-inflammatory lesion counts. However, in Study 18087 for the co-primary endpoint, change in inflammatory lesion count, the comparison of EPIDUO to benzoyl peroxide did not reach statistical significance ($p = 0.068$) for the protocol defined primary analysis method. However, several sensitivity analyses were conducted in which this comparison resulted in p-values less than the nominal $\alpha = 0.05$ level. In the safety assessment of local skin reactions, on average EPIDUO was more irritating than each monad and vehicle, especially within the first week of therapy. The mean intensity of the local skin reaction score for EPIDUO was below a mild rating, and the irritation tended to resolve and reach near baseline levels by week 12.

1.2 Brief Overview of Clinical Studies

In the clinical development of EPIDUO, two adequate and well-controlled twelve week, randomized, double-blind, and parallel group studies (18094 and 18087) were conducted to assess the safety and efficacy of EPIDUO. Each of the two trials were intended to show the contribution of each component of EPIDUO. Study 18094 was a Phase 2 trial and results from this trial were used to power the Phase 3 trial, Study 18087. An additional Phase 3 trial is ongoing, Study 18088, however this trial was not submitted as part of the NDA. In addition a long-term open label study was conducted and submitted to the NDA, Study 18089. Study 18094 enrolled a total of 517 subjects from 36 centers in the U.S. Study 18087 enrolled a total 1668 subjects from 60 centers in the U.S., Puerto Rico, and Canada.

1.3 Statistical Issues and Findings

Study 18094 was a Phase 2 trial. However in the submission of the protocol in SN00 the sponsor proposed that such a trial might be considered as one of two confirmatory Phase 3 trials. In

response to SN00 it was noted that no formal statistical testing is required of Phase 2 trials, however general comments were provided to the sponsor about requirements of the statistical analysis for Phase 3 trials. It should be noted that none of these comments made after the Pre-IND meeting about the statistical analysis were incorporated into a revised protocol for Study 18094. In the review of the protocol for Study 18087, the sponsor and Division did reach agreements on the statistical analysis. For the assessment of efficacy, the agreements reached on the statistical analysis for Study 18087 were also applied to Study 18094. Refer to Section 2.2 for a description of the regulatory history and Section 3.1.2 for details of the statistical analysis agreed upon between the sponsor and the Division.

As Study 18094 enrolled subjects with baseline IGA scores of 'Mild' the assessment of the IGA dichotomized to success was performed using multiple definitions of success in Study 18094. For each of these definitions EPIDUO was consistently statistically superior to each of its monads and vehicle (efficacy results are provided in Table 5 on page 20. In the analysis of the lesion counts, the treatment effects were larger in subjects with higher number of baseline lesions. At baseline it was detected that subjects randomized to EPIDUO had the lowest number of baseline lesions. Thus, as treatment effects were smaller for lower baseline lesion counts such an imbalance in the randomization did not provide a favorable condition to demonstrate efficacy of EPIDUO. Using a main effects model with terms for treatment, baseline lesion count, and site all contrasts comparing EPIDUO to its monads and vehicle were statistically significant at the $\alpha = 0.05$ level. Efficacy results for the lesion counts can be found in Tables 7 (inflammatory lesions), 8 (non-inflammatory lesions), and 9 (total lesions) on pages 23, 26, and 26, respectively.

Study 18087 was powered at over 90% using estimates of treatment effects from Study 18094. Based upon the pre-specified criteria, EPIDUO was superior to each of its monads and vehicle for the co-primary dichotomized IGA endpoint as well as for the change in non-inflammatory lesion endpoint. Efficacy results for these two endpoints are provided in Tables 17 and 19 on pages 34 and 38. For the co-primary endpoint, change in inflammatory lesion count, the comparison of EPIDUO to benzoyl peroxide did not reach statistical significance ($p = 0.068$) for the protocol defined primary analysis method. This analysis was based on ranking the data prior to fitting the ANCOVA model as ranks were used due to the significance of Shapiro-Wilk's normality test. An examination of the ANCOVA model diagnostics did not reveal an added benefit to using the ranks over the untransformed data in which case the analysis based on the untransformed data reached statistical significance ($p = 0.0387$). In addition, a sensitivity analysis using a transformation of the response, end of treatment inflammatory lesion count, with model terms for treatment, site, and baseline count showed more desirable properties in terms of assessing the model diagnostics. This sensitivity analysis also showed a statistically significant difference in inflammatory lesions between EPIDUO and benzoyl peroxide ($p = 0.0199$). Thus, the collective evidence suggests that EPIDUO is statistically superior to each of its monads and vehicle for

all co-primary endpoints.

2 INTRODUCTION

2.1 Product Description

EPIDUO contains Adapalene 0.1% (w/w) and Benzoyl Peroxide 2.5% (w/w) in an aqueous gel vehicle. Both drug substances are approved for the local treatment of acne vulgaris, however no approved combination of the two products has been approved by the Agency. EPIDUO is proposed as a first-line therapy for the topical treatment of acne vulgaris, with a once-daily application [redacted] to the acne affected areas. b(4)

2.2 Regulatory History

The following sections provide a summary of communications about issues which would be influential in the statistical evaluation of the clinical trials made during the review of the IND (IND number was 67,801).

2.2.1 Pre-IND Meeting: 07/28/2003

At the Pre-IND Meeting held on 07/28/2003 the sponsor proposed to conduct a "randomized, multi-center, double-blind, controlled, parallel group study to evaluate safety and efficacy of a fixed-combination of Adapalene 0.1%/Benzoyl Peroxide 2.5% gel compared to the monads Adapalene gel 0.1% and Benzoyl Peroxide 2.5% gel and the fixed-combination gel vehicle in subjects with mild-moderate acne vulgaris." At this time the Division stated the co-primary endpoints should be:

- The success rate based on the Investigator Global Scale (the percentage of patients graded as clear or almost clear) as a static assessment at the efficacy endpoint and not a change from baseline.
- Percent reduction from baseline of facial non-inflammatory and inflammatory lesions.

In Question 1 the sponsor sought concurrence on the plans for the efficacy analysis as listed in Protocol 18094. The following is the response as listed in Biostatistics section of the meeting minutes.

"The sponsor proposed efficacy endpoints based on:

- Percent reduction from baseline in the three facial type counts (non-inflammatory, inflammatory, and total) and

- Success rate defined as the percentage of subjects rated as Clear or Almost Clear on the Investigators Global Assessment Scale

is acceptable (see clinical comments) for establishing efficacy. It should be noted however that as total lesion counts is the sum of the inflammatory and non-inflammatory lesions, for lesion counts endpoint it suffices to consider inflammatory and non-inflammatory lesions. The sponsors definition of the ITT population and the statistical methodology for analysis is acceptable."

In Question 2 the sponsor sought concurrence on the sample size calculation. In response the following comment was provided.

"For Phase 2 trials no formal statistical testing is required. Consequently, Phase 2 trials do not need to be powered. However, the sponsors approach of powering this study is expected to lead to a more reliable estimate of treatment effect compared to that from small trials. Using such an estimate for powering Phase 3 trials would be very helpful in avoiding under-powered Phase 3 trials."

2.2.2 Statistical Review of SN00: 12/15/2003

In SN00 (stamp date: 12/15/2003) the sponsor submitted a revised Phase 2 protocol for Study 18094 which was powered for establishing efficacy and included formal statistical analysis methods. The statistical review also stated that Study 18094 might be used as one of two Phase 3 trials as proposed by the sponsor. The following are some of the comments which were conveyed to the sponsor in the statistical review of SN000.

"The Sponsor should note that usually no formal statistical testing is required, nor the study needs to be powered, as efficacy trend is sufficient for a Phase 2 trial. To establish the efficacy and safety of the drug under investigation, the Division recommends that two Phase 3 trials be conducted. The following are general comments for designing Phase 3 trials.

1. For efficacy claim of acne vulgaris, the combination drug needs to establish superiority to each monad and vehicle in:
   - Change from baseline in inflammatory and non-inflammatory lesion counts as the primary analysis. In addition, analysis of percent change in inflammatory and noninflammatory lesion counts should be submitted as supportive.
   - Percent of patients with success in IGA.

   The definition of success category in IGA should be in agreement with the Division.

2. Sample size calculations should be carried out based on the change in inflammatory and noninflammatory lesion counts as well as IGA. In sample size calculations, the variation in the change of inflammatory lesions and non-inflammatory lesions may be different. The Sponsor should provide detailed information about the expected mean lesion change for each treatment for each lesion type, so the study is not under-powered.
3. In addition to the LOCF method for imputing missing data, a sensitivity analysis should be proposed in the protocol to ensure that efficacy results are not driven by the imputation method used.
4. The sponsor proposed several secondary efficacy endpoints over multiple time points. Secondary efficacy endpoints should be clinically relevant and in agreement with the Division. A multiplicity adjustment would be needed if the number of secondary efficacy endpoints intended for labeling were large.
5. The sponsor should plan to enroll a minimum of 10 patients per arm per center for each study to avoid small frequency counts in the analyses."

2.2.3 Statistical Review of SN08: 06/14/2004

In SN08 (stamp date: 06/14/2004) the sponsor submitted a revised protocol for Study 18094. One change made in Protocol 18094 (amendment #1) was to increase the number of sites participating in the study. The original plan was to enroll about 490 patients from 20-30 U.S. study sites. In amendment #1 the number of sites was increased to 30-40 U.S. sites. In response the following is the comment from the statistical review team.

> "To reduce the chance of having small cell frequency in the efficacy analysis, the protocol should be designed to ensure a sufficient enrollment per arm per center. As there is an unequal treatment allocation (2:2:2:1), the sponsor should enroll a minimum of 4-5 patients per site for the smallest arm (i.e., vehicle)."

It should be noted that the statistical comments provided in the review of SN00 were not incorporated into amendment #1 of Protocol 18094.

2.2.4 End of Phase 2 Meeting: 12/12/2005

An End of Phase 2 Meeting was held on 12/12/2005. At this time the sponsor requested the Division's concurrence on using Study 18094 as one of two efficacy and safety studies to support filing the NDA. In response the clinical and biostatistics review teams had the following comment.

"Whether the completed Study [18094] can be used to establish the efficacy claim for Adapalene/Benzoyl Peroxide Topical Gel is a review issue which will depend on the study design, statistical method of analysis, and the efficacy findings. In general, the agency requires efficacy established based on two well-designed independent Phase 3 trials. The Division stated that [S]tudy 18094 was a phase 2 trial and the study synopsis stated that Study unblinded as prospectively defined in the protocol. It is not clear when the unblinding was done. In addition, the study was powered at 80% to detect a 15% difference in success rate and percent change in lesion counts. It should be noted that the summary of efficacy results was 10% difference in the success rate with the IGA, 4 lesions for change in inflammatory lesions and 6 lesions for change of noninflammatory lesions, yet all the reported p-values were approximately 0.001."

The sponsor also sought concurrence on the sample size justification for Study 18087. In response the Division stated, "the study should be powered for the two co-primary endpoints:

1. The success rate at week 12. Success is defined as 0 (clear) or 1 (almost clear) or alternatively success could defined as improvement of two grades from the baseline score on the IGA.

2. Change in inflammatory, noninflammatory and total lesion counts."

In addition the sponsor sought the Division's concurrence on the adequacy of the statistical analysis for Study 18087. The following is a summary of some of the Division's comments about the statistical analysis plan for Study 18087.

1. "Use of the ITT population as the primary population and the PP as supportive is acceptable. The ITT population should be defined as all patients randomized and dispensed medication, whether or not they have any post-baseline assessments. Also, the sponsor should note that a list of criteria excluding subjects from PP analysis population should be defined in the protocol.
2. For imputation of missing data, the Sponsor's proposed approach of LOCF along with sensitivity analyses by imputing missing data as success in one analysis and failure in another analysis should be acceptable for the IGA. For imputation of missing data for the other co-primary endpoint, lesion counts, LOCF approach might be used along with a sensitivity analysis by imputing the missing data as the mean (median) lesion count in each arm.
3. It is not clear what the sponsor meant by lesion counts (ranked). Analysis of lesion counts should be done on the original scale if the underlying assumptions

of the statistical methodology, such as normality, hold. The protocol should pre-specify an approach for testing such assumptions and propose an alternative approach for the analysis, (e.g. nonparametric analysis) if assumptions are not met.

4. In addition to testing normality of the data, ANCOVA model assumes a linear relationship between the covariate and the mean response, with equal slopes for each treatment. The protocol should test for equality of slopes of the different treatment regression lines (test for parallel slopes)."

2.2.5 Statistical Review of SN22: 02/06/2006

Protocol 18087 was submitted for Special Protocol Assessment on 02/06/2006. In the review of the protocol there were several outstanding statistical issues. The following is a brief summary of the outstanding issues at the time.

1. Reiteration about the acceptability of Study 18094 as one of two Phase 3 trials needed to establish efficacy.

2. Encouragement to further reduce the number of centers participating in this study to reduce the chance of having sparse tables and its impact on the analysis.

3. Guidance on sensitivity analysis to assess the impact of missing data on the study outcome.

4. Guidance on how to assess lesion counts with an ANCOVA model.

5. Statement that the quality of life questionnaire [redacted] b(4)

2.2.6 Statistical Review of SN26: 02/06/2006

A protocol amendment was made to Protocol 18087 to address both clinical and statistical comments provided in the SPA review. At this point the main statistical issue was that the protocol did not include a multiplicity adjustment for the assessment of secondary endpoints. Subsequent to the review of SN26 no further statistical reviews of Protocol 18087 were required as the sponsor and Division were in agreement with the statistical analysis methods proposed. Details of the agreed upon statistical methodology are described in Section 3.1.2.

2.2.7 Pre-NDA Meeting: 11/28/2007

A Pre-NDA Meeting was held on 11/28/2007. During this meeting the sponsor clarified that the data were not unblinded in Study 18094. Overall the discussion was based on the content of the NDA and what the sponsor should submit for sufficient review by the Division.

2.3 Clinical Trial Overview

In the clinical development of EPIDUO, two adequate and well-controlled twelve week, randomized, double-blind, and parallel group studies (SRE.18094 and SRE.18087) were conducted to assess the safety and efficacy of EPIDUO. Each of the two trials were intended to show the contribution of each component of EPIDUO . Study 18094 was a Phase 2 trial and results from this trial were used to power the Phase 3 trial, Study 18087. An additional Phase 3 trial is ongoing, Study 18088, however this trial was not submitted as part of the NDA. In addition a long-term open label study was conducted and submitted to the NDA, Study 18089. Table 1 provides a brief overview of each trial.

Table 1: Efficacy and Safety Studies Overview

Study	Development Objective	Drug Products	Number Subjects	Date†
SRE.18094 (Study 18094)	Phase 2 Superiority	EPIDUO	149	02/2004 – 12/2004
		Adapalene	148	
		Benzoyl Peroxide	149	
		Vehicle	71	
SRE.18087 (Study 18087)	Phase 3 Superiority	EPIDUO	415	06/2006 – 07/2007
		Adapalene	420	
		Benzoyl Peroxide	415	
		Vehicle	418	
SRE.18089 (Study 18089)	Phase 3 Long-term Safety	EPIDUO	452	02/2004 – 05/2005
			-	

† Dates correspond to the start and end of the study.

The review of efficacy is based on the two vehicle-controlled trials, Study 18094 and Study 18087. The review of of short-term safety is based on Study 18094 and 18087. Assessment of long-term safety is based on Study 18089.

2.4 Data Sources

In the review of the electronic data sets submitted to the NDA, several deficiencies were identified. Consequently the Division contacted the sponsor and requested the sponsor resubmit new electronic data sets to assess the efficacy of EPIDUO. Below is the issue identified and the request for information.

Issue: Efficacy results as presented in the study report for Studies 18087 and 18094 appear to rely on the use of the derived analysis visit variable XVISIT. However,

the sponsor has not included details of how this analysis visit variable was defined. Further, this analysis visit is not included in all data sets which does not allow merging of data sets by unique subject ID and visit. Consequently, this does not allow for a thorough review of the data submitted in the NDA.

Request for Information: The sponsor is requested to submit new electronic data sets which include a consistent definition of the analysis visit allowing for merging of data sets by unique subject ID AND visit. Any derived analysis visit should also include a detailed description of how such a variable was derived. In addition, it is requested that the date of the visit also be included in the electronic data. The sponsor should note that results as presented in the study reports should be reproducible using the submitted data sets.

In addition to the above request, the sponsor was provided with an example to better elicit the preferred structure of the efficacy data sets. On 04/22/2008 the sponsor submitted two analysis data sets for Studies 18094 and 18087; one corresponded to IGA assessment, the other to assessment of lesion counts. The revised analysis data sets as requested by the Division are located at //Cdsesub1/evsprod/NDA022320/0002/m5/datasets/rd-06-sre-180xx/analysis.

3 Statistical Evaluation

3.1 Evaluation of Efficacy

3.1.1 Studies

Study 18094 was a Phase 2 trial. The protocol for this study was reviewed by the statistical review team for a Pre-IND meeting held on 07/28/2003 as well as in submissions SN000 (stamp date: 12/15/2003) and SN008 (stamp date: 06/14/2004). Refer to Section 2.2 for a detailed account of the comments provided to the sponsor in the review of Protocol 18094.

Study 18087 was one of two Phase 3 trials. The Division provided several comment on the design of the trial and statistical methodologies including an End of Phase 2 meeting on 12/12/2005 in addition to statistical comments provided in the review of SN022 (stamp date: 02/06/2006) and SN026 (stamp date: 06/07/2006).

Study 18088 is ongoing and not included in the NDA. The design of Study 18088 is similar to the design of Study 18087 and was subject to review in SN033 (stamp date: 08/24/2006). Further statistical issues relevant to both Phase 3 trials were also provided in the statistical reviews of SN044 (stamp date: 01/30/2007) and SN054 (stamp date: 04/20/2007).

3.1.2 Statistical Methodology

As described in Section 2.2 several comments were provided in the review of the protocol for Study 18094 which were not incorporated into the statistical analysis section of the protocol. In the review of the protocol for Study 18087 the Division and the sponsor reached agreements on the statistical analysis of the primary endpoints. As the statistical analysis details are more well-defined for Study 18087 and the endpoints are in agreement with the Division, these agreed upon statistical methodologies are applied to both studies submitted to the NDA. Thus, the statistical methodologies described below correspond to those included in the protocol for Study 18087 and not those included in the protocol for Study 18094, any deviations from protocol defined methods are documented.

3.1.2.1 Populations The primary analysis population is defined as the intent-to-treat (ITT) population which includes all subjects who were randomized and dispensed medication. The per protocol (PP) population which excludes subjects with major protocol violations is planned as a supportive analysis to the primary analysis on the ITT population.

3.1.2.2 Statistical Analysis All comparisons of EPIDUO to its monads and vehicle for the co-primary endpoints will be tested at the two-sided $\alpha = 0.05$ significance level. Small centers will be pooled prior to analysis which combines the largest center with the smallest center. These pooled centers will be referred to as "analysis centers" in the statistical analyses. The trial will meet efficacy criteria if all primary analyses are shown to be statistically significant at the two-sided $\alpha = 0.05$ level.

3.1.2.2.1 Investigator Global Assessment For the analysis on the percent of subjects with an IGA success, the protocol listed CMH stratified by "analysis center" as the primary method of statistical analysis. The primary method of data imputation is LOCF with two sensitivity analyses listed as (a) impute all missing week 12 data as failures and (b) impute all missing week 12 data as successes.

3.1.2.2.2 Lesion Counts Change in inflammatory and non-inflammatory lesion counts from baseline to week 12 (LOCF, ITT) will be analyzed by two-way ANCOVA model including baseline lesion count as a covariate, treatment, analysis center, and treatment-by-baseline as factors. Treatment-by-center interaction will be examined and will be included in the model if qualitative interaction is detected. The normality assumption will be tested at the 0.05 significance level using Shapiro-Wilks tests on the residuals from the ANCOVA model.

If the normality assumption is met, the equality of slopes will be tested to determine whether equal or unequal slopes (i.e. treatment-by-baseline) should be included in the final model. If

the treatment effect is different for different baseline lesion counts, further analysis by baseline lesion counts using the quartiles will be provided. If the normality assumption is not met, the ranked change in lesion count will be analyzed by two-way ANCOVA model using ranked baseline count as a covariate, and treatment and analysis center as factors. Note that in the end only a single model was fit for each lesion count and contrasts were used to test for treatment effect differences between EPIDUO and each of its monad and vehicle which was not explicitly stated in the protocol.

Reviewer Comment: Note that in all the models fit, the Shapiro-Wilk's test on the residuals were highly significant ($p < 0.0001$) when using the original unranked data as well as when the lesion counts were ranked. In addition, the diagnostic examination of the ranked data showed no improved properties over the unranked data. While the sponsor's study report depicts efficacy results using ranks, the efficacy results reported in this review are primarily on the unranked data as this aids in interpreting the model. Due to possible violations of the ANCOVA model assumptions a sensitivity analysis is conducted in this review which explores the ANCOVA model assumptions. The methodology of the sensitivity analysis uses end of treatment count as the response with terms for treatment, analysis center, and baseline count.

The primary method for data imputation of the change in lesion counts is LOCF. As sensitivity analyses missing week 12 lesion counts will be imputed using two strategies which is consistent with the missing data sensitivity analysis of IGA: (a) impute the median change in lesion count from the IGA 'Failures' for each treatment group and (b) impute the median change in lesion count from the IGA 'Success' for each treatment group.

3.1.3 Efficacy: Study 18094

3.1.3.1 Study Design Study 18094 was a Phase 2 multi-center, randomized, double-blind, active- and vehicle-controlled study to evaluate the safety and efficacy of EPIDUO compared with adapalene, benzoyl peroxide, and vehicle in the treatment of acne vulgaris when applied once daily. Male and female subjects, age 12 years or older, with 20 to 50 inflammatory lesions and 30 to 100 non-inflammatory lesions on the face were randomized at the baseline visit, instructed to apply study medication once daily in the evening to treat the affected areas for up to 12 weeks. Note that inclusion criteria did not require a specific baseline IGA score for study enrollment though an IGA score was recorded for baseline which ranged from mild to severe[1]. 517 subjects were recruited and randomized in 2:2:2:1 ratio to EPIDUO, adapalene, benzoyl peroxide, and vehicle, respectively at 36 centers in the U.S.

The treatment period consisted of 12 weeks with 6 study visits occurring at baseline, week

[1]At the Pre-IND Meeting held on July 28, 2003 the Division stated the use of lesion counts and other inclusion/exclusion criteria were acceptable.

1, week 2, week 4, week 8, and week 12. The primary timepoint for efficacy evaluation is at the week 12 visit which was planned to occur within a 7 day treatment window. The protocol stated efficacy objective of the trial was to demonstrate the clinical efficacy of EPIDUO compared with adapalene, benzoyl peroxide, and vehicle as assessed by the primary efficacy criteria of IGA success ("Clear" or "Almost Clear") and percent reduction from baseline of facial non-inflammatory, inflammatory lesion counts, and total lesion counts. In addition, the trial was designed to assess the safety and tolerability profile of EPIDUO as assessed by the local tolerability parameters (erythema, scaling, dryness and stinging/burning) and adverse event reporting.

3.1.3.2 Endpoints The following comment was provided in the statistical review of the protocol for Study 18094 which was submitted in SN000 (stamp date: 12/15/2003).

> "For an efficacy claim of acne vulgaris, the combination drug needs to establish superiority to each monad and vehicle in:
>
> - Change from baseline in inflammatory and non-inflammatory lesion counts as the primary analysis. In addition, analysis of percent change in inflammatory and noninflammatory lesion counts should be submitted as supportive.
> - Percent of patients with success in IGA.
>
> The definition of success category in IGA should be in agreement with the Division."

Study 18094 was initiated without obtaining an agreement on the definition of IGA success and the study enrolled subjects with baseline IGA scores of "Mild". While the protocol defines success as "Clear" or "Almost Clear", multiple definitions of treatment success on the IGA scale are provided in this statistical evaluation of the IGA scale for Study 18094. Also note that at this time the Division preferred to use absolute change in lesion counts rather than percent change.

3.1.3.3 Patient Disposition and Baseline Characteristics

3.1.3.3.1 Patient Disposition A total of 45 out of 517 (8.7%) subjects discontinued from Study 18094; reason recorded for subject discontinuation is provided in Table 2. The most prevalent reason for subject discontinuation was due to subject request which accounted for 28 of the 45 subjects who discontinued. The overall percent of subjects that completed the trial was greater than 85% for all treatment arms.

Table 2: Subject Disposition (Study 18094)

	EPIDUO ($N = 149$)	Adapalene ($N = 148$)	Benzoyl Peroxide ($N = 149$)	Vehicle ($N = 71$)
Complete Trial	139 (93.3)	131 (88.5)	139 (93.3)	63 (88.7)
Discontinued	10 (6.7)	17 (11.5)	10 (6.7)	8 (11.3)
Lack of Efficacy	0 (0)	1 (0.7)	0 (0)	0 (0)
Adverse Event	1 (0.7)	1 (0.7)	0 (0)	0 (0)
Subject Request	4 (2.7)	10 (6.8)	7 (4.7)	7 (9.9)
Protocol Violation	2 (1.3)	1 (0.7)	0 (0)	0 (0)
Lost to Follow-up	2 (1.3)	3 (2)	3 (2)	1 (1.4)
Other	1 (0.7)	1 (0.7)	0 (0)	0 (0)

Source: Study Report Table 5; results reproduced by reviewer

3.1.3.3.2 Baseline Demographic Factors The baseline demographics for age, gender, and race are provided in Table 31 in the Appendix. Overall, the mean age of subjects was 16 years old; more than 70% of subjects were identified as Caucasian; and 40% of subjects were female. There was no imbalance of any of the demographic factors between treatment arms.

3.1.3.3.3 Baseline Prognostic Factors This exploratory analysis examines various baseline characteristics of acne vulgaris that may be related to disease severity. Ideally, we want to see that randomization to treatment for the four treatment arms resulted in similar baseline values of the prognostic characteristics as large disparities could have an effect on safety and efficacy claims. The prognostic characteristics examined include the following.

- IGA: 5-point scale with 0 = clear and 4 = severe.
- Burning: 4-point scale with 0 = none to 3 = severe.
- Dryness: 4-point scale with 0 = none to 3 = severe.
- Erythema: 4-point scale with 0 = none to 3 = severe.
- Scaling: 4-point scale with 0 = none to 3 = severe.

Also examined are the baseline values of lesion counts including total lesion count, inflammatory lesion count, and non-inflammatory lesion count.

Summary statistics are provided in Table 3. The majority of subjects had baseline IGA scores of 'Moderate' which varied from 75% of subjects randomized to adapalene to 85% of subjects randomized to benzoyl peroxide. In terms of the baseline lesion counts, subjects randomized to EPIDUO had lower median counts for all lesion counts than the other treatment arms. Looking

at an empirical cumulative distribution of the lesion counts by treatment group it shows that the proportion of subjects with any given lesion count is always smallest for subjects randomized to EPIDUO (empirical cumulative distribution plot provided in the Appendix, Figure 19). A density plot of the baseline lesion counts (not shown) confirms that the the baseline distributions of lesion counts are similar but the there is a shift to the left in the density of baseline lesion counts for EPIDUO.

Table 3: Baseline Prognostic Factors by Treatment (Study 18094)

	EPIDUO $N = 149$	Adapalene $N = 148$	Benzoyl Peroxide $N = 149$	Vehicle $N = 71$
IGA : Mild	17% (25)	19% (28)	10% (15)	18% (13)
Moderate	80% (119)	75% (111)	85% (127)	80% (57)
Severe	3% (5)	6% (9)	5% (7)	1% (1)
Inflammatory	6.00 11.00 20.00	8.75 15.00 25.25	8.00 16.00 26.00	12.50 19.00 28.00
Non-Inflammatory	13.0 22.0 37.0	18.0 32.0 47.0	15.0 27.0 45.0	20.0 33.0 49.5
Total	22.00 36.00 54.00	27.75 52.00 72.00	29.00 47.00 74.00	36.50 49.00 85.00
Burning : None	97% (144)	100% (148)	97% (145)	99% (70)
Mild	3% (5)	0% (0)	3% (4)	1% (1)
Moderate	0% (0)	0% (0)	0% (0)	0% (0)
Severe	0% (0)	0% (0)	0% (0)	0% (0)
Dryness : None	89% (132)	92% (136)	90% (134)	93% (66)
Mild	11% (17)	7% (11)	10% (15)	6% (4)
Moderate	0% (0)	1% (1)	0% (0)	1% (1)
Severe	0% (0)	0% (0)	0% (0)	0% (0)
Erythema : None	64% (95)	61% (91)	63% (94)	63% (45)
Mild	31% (46)	36% (54)	32% (48)	25% (18)
Moderate	5% (8)	2% (3)	5% (7)	11% (8)
Severe	0% (0)	0% (0)	0% (0)	0% (0)
Scaling : None	92% (137)	93% (138)	92% (137)	93% (66)
Mild	8% (12)	7% (10)	8% (12)	7% (5)
Moderate	0% (0)	0% (0)	0% (0)	0% (0)
Severe	0% (0)	0% (0)	0% (0)	0% (0)

a b c represent the lower quartile a, the median b, and the upper quartile c for continuous variables. Numbers after percents are frequencies.

Source: Reviewers analysis.

The local skin reactions: burning, dryness, erythema, and scaling were ranked from none (0) to severe (5). At baseline over 90% of subjects scored dryness, burning, and scaling as 'none'. Approximately 30% of subjects scored erythema as 'Mild' and even a small percentage had scores of 'Moderate'. Overall the distributions were balanced across treatment arms for local

skin reactions.

3.1.3.4 Populations Analyzed The Intent-to-Treat population (ITT) includes all subjects randomized and dispensed study medication. Subjects with major protocol deviations were excluded from the Per-Protocol (PP) population. Table 4 depicts the number of subjects included in the two analysis populations for each of the studies.

Table 4: Summary of Data Sets Analyzed

	EPIDUO	Adapalene	BPO	Vehicle
ITT Population	149	148	149	71
PP Population	125	116	129	51

Source: Sponsor's Study Report Table 8; results reproduced by reviewer.

3.1.3.5 Primary Endpoint Results (ITT-LOCF) The section on efficacy is broken down into two sections. The first examines efficacy results according to the investigator global assessment whereas the second examines efficacy results based upon change in lesion counts.

3.1.3.5.1 Investigator Global Assessment In the review of the protocol for Study 18094 it was communicated to the sponsor that the definition of success using the IGA should be in agreement with the Division. The protocol lists the definition of treatment of success based on the IGA as subjects with an IGA score of 'Clear' or 'Almost Clear'. As the enrollment criteria allowed subjects to enroll with a baseline IGA score of 'Mild' these subjects can have a one grade improvement to 'Almost Clear' and be defined as success based on such a definition of IGA success. Typically the Division requests subjects enrolling with 'Mild' IGA scores to reach 'Clear' to be considered a treatment success. As such, the review considers multiple definitions of treatment success which are listed below.

- 'Clear' or 'Almost Clear': Subjects with an IGA score of 'Clear' or 'Almost Clear' at the end of treatment visit are considered a treatment success.
- Two Grade Improvement: Subjects who have a two grade improvement from their baseline IGA score are considered a treatment success.
- Intersecting Definition: To be defined as a treatment success subjects must have a two grade improvement from the baseline IGA score AND reach a score of 'Clear' or 'Almost Clear'.

Using the above definitions of treatment success based on the IGA scale, Table 5 depicts the efficacy results for Study 18094. This table shows that for each definition of IGA success, EPIDUO was statistically superior to each monad and the vehicle at the $\alpha = 0.05$ significance level when using CMH stratified by analysis center.

Table 5: Investigator Global Results (ITT-LOCF)

	EPIDUO ($N = 149$)	Adapalene ($N = 148$)	BPO ($N = 149$)	Vehicle ($N = 71$)
Clear or Almost Clear†				
Success (%)	41 (27.5)	23 (15.5)	23 (15.4)	7 (9.9)
p-value	-	0.0079	0.0034	0.0015
Two Grade Improvement*				
Success (%)	33 (22.1)	19 (12.8)	18 (12.1)	4 (5.6)
p-value	-	0.0309	0.0056	0.0016
Intersecting Definition*				
Success (%)	32 (21.5)	18 (12.2)	18 (12.1)	4 (5.6)
p-value	-	0.0291	0.0088	0.0023

† Source: Sponsor's Study Report Table 12; results reproduced by reviewer.

* Source: Reviewer Analysis.

As a sensitivity analysis, mild subjects were excluded and success was defined as 'Clear' or 'Almost Clear' which mimics the population and IGA success for Study 18087. Efficacy results are depicted in Table 6. Response rates in this sensitivity analysis are consistent with those in the primary analysis with treatment effects comparing EPIDUO to each of its monads and vehicle being at least 10%. Despite the smaller sample sizes, all comparisons reached the nominal $\alpha = 0.05$ level.

Table 6: Investigator Global Results (ITT-LOCF-IGA at Least Moderate at Baseline)

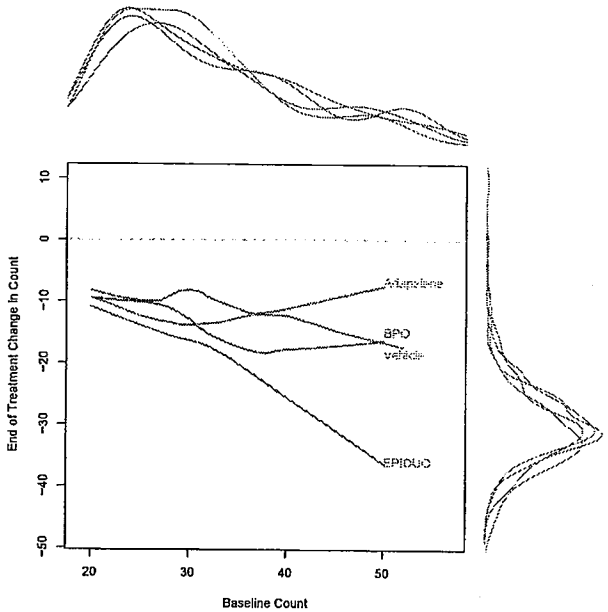
	EPIDUO	Adapalene	BPO	Vehicle
	($N = 124$)	($N = 120$)	($N = 134$)	($N = 58$)
Success (%)	31 (25.0)	18 (15.0)	18 (13.4)	4 (6.9)
p-value†	-	0.0336	0.0067	0.0023

† Tests are CMH stratified by site.

Source: Reviewer Analysis.

3.1.3.5.2 Change in Lesion Counts Figure 1 depicts a loess nonparametric regression line[1] for the inflammatory lesions at baseline and the change in inflammatory lesion count along with univariate density estimates of the distributions for each treatment group[2]. The loess lines show that for subjects with a low baseline inflammatory lesion count there is only a slight separation between EPIDUO and the monads. However, when the baseline number of inflammatory lesions counts increases it can be seen that EPIDUO has a larger decrease in the number of end of treatment inflammatory lesions than the monads and vehicle. Overall, the plot shows that the treatment profiles over the range of baseline counts are not parallel among treatments suggesting a possible interaction between baseline count and treatment.

Figure 1: Inflammatory Lesion Counts (Study 18094)



In the analysis of change in inflammatory lesion counts there was no significant treatment by analysis center interaction and the fitted model with estimated regression parameters is shown in the equation on the following page. As the model shows there is a significant baseline by treatment interaction.

[2]Note that the loess fits are depicted rather than the raw data points to provide a visual depiction of the nonparametric relationship between baseline and end of treatment lesion counts.

$$\mathrm{E}(\Delta_{\mathrm{INF}}) = X\beta, \quad \text{where}$$

$$\begin{aligned} X\hat{\beta} = \; & -0.3677 \\ & +8.376\{\mathrm{Site}_{102}\} + 4.020\{\mathrm{Site}_{103}\} + 8.544\{\mathrm{Site}_{104}\} + 10.215\{\mathrm{Site}_{105}\} + 12.715\{\mathrm{Site}_{106}\} \\ & +12.078\{\mathrm{Site}_{107}\} + 6.801\{\mathrm{Site}_{108}\} + 10.026\{\mathrm{Site}_{109}\} + 9.356\{\mathrm{Site}_{110}\} + 4.751\{\mathrm{Site}_{111}\} \\ & +8.606\{\mathrm{Site}_{112}\} + 3.765\{\mathrm{Site}_{113}\} - 0.784\,\mathrm{Base}_{\mathrm{INF}} - 10.90\{\mathrm{BPO}\} - 14.10\{\mathrm{Adapalene}\} \\ & -9.64\{\mathrm{Vehicle}\} + \mathrm{Base}_{\mathrm{INF}}[0.5619\,\{\mathrm{BPO}\} + 0.6298\,\{\mathrm{Adapalene}\} + 0.5477\,\{\mathrm{Vehicle}\}] \end{aligned}$$

and $\{c\} = 1$ if in group c, 0 otherwise.

Essentially, the estimated parameters for a given site just alter the estimate of the intercept of the predicted change in lesion counts. As such, to get an idea of the predicted change in inflammatory lesion counts for a given baseline inflammatory lesion count, Figure 2 depicts the predicted change in inflammatory lesion counts for each treatment in analysis center 108[3]. This shows that for subjects with a baseline inflammatory lesion count of 20, the predicted reduction in inflammatory lesion counts is highest in subjects treated with adapalene. However, as the number of baseline inflammatory lesions increase, the model predicts the greatest reduction in subjects treated with EPIDUO.

A summary of the change as well as the percent reduction in inflammatory lesion counts is provided in Table 7. The table also provides p-values for testing the main treatment effect for a model with the main effects only not the model as described above with the interaction term. While there was a significant baseline by treatment interaction, a model with only main effects shows a EPIDUO is statistically superior to each of its monads and vehicle.

[3]As the estimated regression model suggests, different centers just alter the intercept of the regression lines, and center 108 is only used for illustrative purposes.

Figure 2: ANCOVA Fitted Inflammatory Lesion Counts (Study 18094)

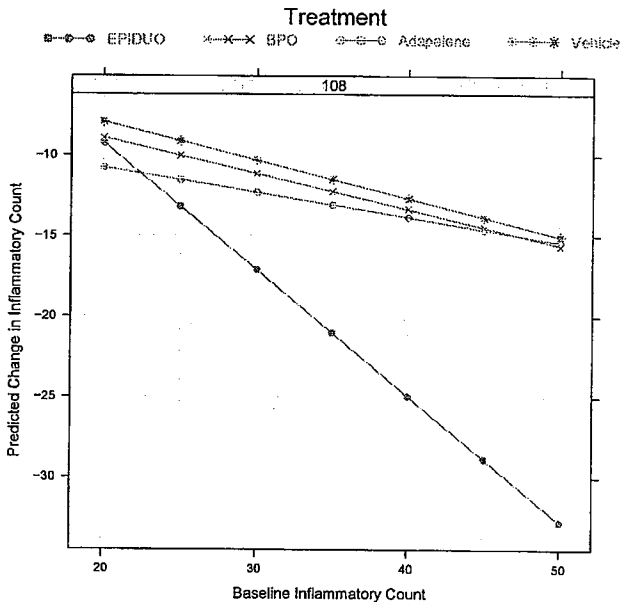


Table 7: Change in Inflammatory Lesion Counts (ITT-LOCF): 18094

	EPIDUO ($N = 149$)	Adapalene ($N = 148$)	BPO ($N = 149$)	Vehicle ($N = 71$)
Mean Change	-16.0	-11.4	-10.5	-9.5
Mean Percent Change	-52.4	-39.9	-35.8	-31.8
p-value†	-	< 0.001	< 0.001	< 0.001
p-value‡	-	.0012	< 0.001	< 0.001

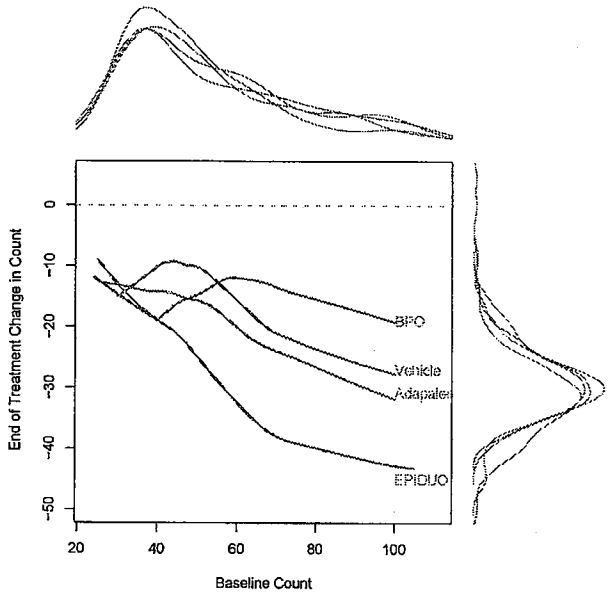
† P-values are based on the ANCOVA model on rank data of changes from baseline lesion counts, including rank data of baseline lesion count as a covariate, treatment and center as main effects. Source: Sponsor's Table 22 in Module 2.

‡ Source: Reviewer's analysis using an ANCOVA model with main effects only on the unranked data.

Figure 3 depicts a smoothed regression line for the non-inflammatory lesions at baseline and the change in non-inflammatory lesion count along with univariate density estimates of the distributions for each treatment group. The smoothed regression lines show that for subjects with a low baseline non-inflammatory lesion count there is little separation between the treatment groups implying a minimal treatment effect. However, when the baseline number of non-inflammatory lesions counts increases it can be seen that EPIDUO has a higher change in

the number of end of treatment non-inflammatory lesions than the monads and vehicle. Overall, the plot shows that the treatment profiles over the range of baseline counts are not parallel among treatments suggesting a possible interaction between baseline count and treatment.

Figure 3: Non-Inflammatory Lesion Counts (Study 18094)



In the analysis of change in non-inflammatory lesion counts there was no significant treatment by analysis center interaction and the fitted model with estimated regression parameters is shown in the equation on the following page. As the model shows there is a significant baseline by treatment interaction.

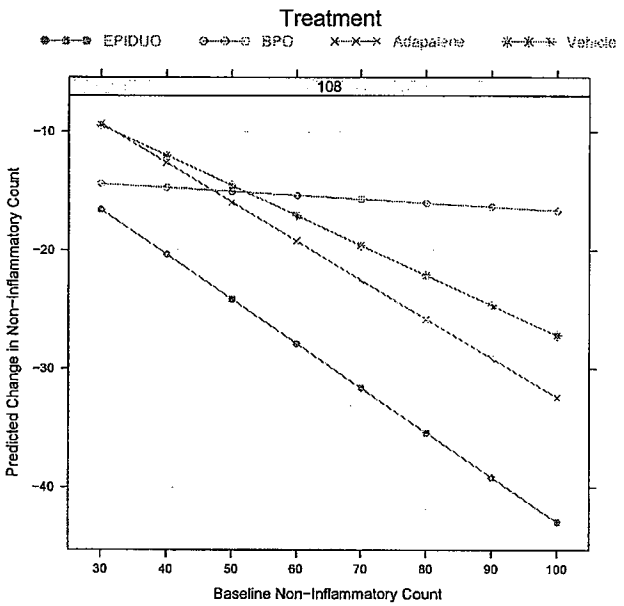
$$\mathrm{E}(\Delta_{\mathrm{NonINF}}) = X\beta, \quad \text{where}$$

$$\begin{aligned} X\hat{\beta} = \; & -17.06 \\ & +12.916\{\mathrm{Site}_{102}\} + 13.279\{\mathrm{Site}_{103}\} + 13.014\{\mathrm{Site}_{104}\} + 16.483\{\mathrm{Site}_{105}\} + 15.653\{\mathrm{Site}_{106}\} \\ & +10.247\{\mathrm{Site}_{107}\} + 11.798\{\mathrm{Site}_{108}\} + 17.698\{\mathrm{Site}_{109}\} + 23.038\{\mathrm{Site}_{110}\} + 9.435\{\mathrm{Site}_{111}\} \\ & +9.896\{\mathrm{Site}_{112}\} + 15.127\{\mathrm{Site}_{113}\} - 0.3768\,\mathrm{Base}_{\mathrm{NonINF}} \\ & -8.136\{\mathrm{BPO}\} + 5.839\{\mathrm{Adapalene}\} + 3.386\{\mathrm{Vehicle}\} \\ & +\mathrm{Base}_{\mathrm{NonINF}}[0.34423\,\{\mathrm{BPO}\} + 0.04698\,\{\mathrm{Adapalene}\} + 0.12392\,\{\mathrm{Vehicle}\}] \end{aligned}$$

and $\{c\} = 1$ if subject is in group c, 0 otherwise.

The predicted change in non-inflammatory lesion counts for a given baseline non-inflammatory lesion count is depicted in Figure 4 for each treatment in analysis center 108. The figure suggests that the predicted change in non-inflammatory lesions for benzoyl peroxide is the cause of the significant interaction term. Overall, the predicted change in non-inflammatory lesions is greatest for subjects randomized to EPIDUO.

Figure 4: ANCOVA Fitted Non-Inflammatory Lesion Counts (Study 18094)



A summary of the change as well as the percent reduction in non-inflammatory lesion counts is provided in Table 8. The table also provides p-values for testing the main treatment effect for a model with the main effects only not the model as described above with the interaction term. While there was a significant baseline by treatment interaction, a model with only main effects shows a EPIDUO is statistically superior to each of its monads and vehicle.

Table 8: Change in Non-Inflammatory Lesion Counts (ITT-LOCF): 18094

	EPIDUO ($N = 149$)	Adapalene ($N = 148$)	BPO ($N = 149$)	Vehicle ($N = 71$)
Mean Change	-23.4	-15.2	-13.7	-13.2
Mean Percent Change	-45.9	-29.6	-32.2	-27.8
p-value†	-	< 0.001	< 0.001	< 0.001
p-value‡	-	0.0001	.0001	0.0003

† P-values are based on the ANCOVA model on rank data of changes from baseline lesion counts, including rank data of baseline lesion count as a covariate, treatment and center as main effects. Source: Sponsor's Table 22 in Module 2.

‡ Source: Reviewer's analysis using an ANCOVA model with main effects only on the unranked data.

A summary of the change as well as the percent reduction in total lesion counts is provided in Table 9. The table also provides p-values for testing the main treatment effect for a model with the main effects only. In terms of change in number of total lesions, EPIDUO was statistically superior to each of its monads and vehicle.

Table 9: Change in Total Lesion Counts (ITT-LOCF): 18094

	EPIDUO ($N = 149$)	Adapalene ($N = 148$)	BPO ($N = 149$)	Vehicle ($N = 71$)
Mean Change	-39.3	-26.5	-24.1	-22.6
Mean Percent Change	-48.5	-34.0	-33.3	-29.7
p-value†	-	< 0.001	< 0.001	< 0.001
p-value‡	-	< 0.001	< 0.001	< 0.001

† P-values are based on the ANCOVA model on rank data of changes from baseline lesion counts, including rank data of baseline lesion count as a covariate, treatment and center as main effects. Source: Sponsor's Table 22 in Module 2.

‡ Source: Reviewer's analysis using an ANCOVA model with main effects only on the unranked data.

3.1.3.6 Primary Endpoint Results (PP-LOCF)

3.1.3.6.1 Investigator Global Assessment Similar to the efficacy assessment on the ITT population, efficacy results for multiple definitions of treatment success on the IGA scale

are assessed as subjects were allowed to enroll with a baseline IGA score of 'Mild' in Study 18094 . Results for the PP population are presented in Table 10.

Table 10: Investigator Global Results (PP-LOCF): 18094

	EPIDUO ($N = 125$)	Adapalene ($N = 116$)	BPO ($N = 129$)	Vehicle ($N = 51$)
Clear or Almost Clear†				
Success (%)	37 (29.6)	21 (18.1)	21 (16.3)	7 (13.7)
p-value	-	0.0188	<.001	0.0105
Two Grade Improvement*				
Success (%)	30 (24)	18 (15.5)	16 (12.4)	4 (7.8)
p-value	-	0.0539	0.0013	0.0044
Intersecting Definition*				
Success (%)	29 (23.2)	17 (14.7)	16 (12.4)	4 (7.8)
p-value	-	0.0524	0.0023	0.0065

† Source: Sponsor's Study Report Table EFF 1.3; results reproduced by reviewer.

* Source: Reviewer Analysis.

3.1.3.6.2 Change in Lesion Counts For the assessment of absolute change in lesion count in the PP population, the main effects model is fit with absolute change in lesion counts as the response and terms for treatment, analysis center, and baseline lesion count (see equation 1). Note that the sponsor reported efficacy results using ranks on the lesion counts, but as previously discussed this does not improve the normality of the residuals. As such efficacy results provided below are based upon data that is not ranked.

Efficacy results for the change in inflammatory, non-inflammatory, and total lesion counts are provided in Table 11 for the PP population. The point estimates for change and percent change show more reduction in the baseline lesions in the PP population than the ITT population. Otherwise, p-values are consistent with the ITT population, though different models were used to assess treatment effects.

3.1.3.7 Missing Data Sensitivity Analysis As a sensitivity analysis to the primary method of data imputation as LOCF, two alternate imputation approaches were provided in the protocol and agreed up with the Division. For the IGA the alternate imputation approached are: (a) impute all missing week 12 data as failures and (b) impute all missing week 12 data as successes. For the change in lesion counts the alternate imputation approaches are: (a) impute the median change in lesion count from the IGA 'Failures' for each treatment group and (b) impute the median change in lesion count from the IGA 'Success' for each treatment group.

Table 11: Change in Lesion Counts (PP-LOCF): 18094

	EPIDUO ($N = 125$)	Adapalene ($N = 116$)	BPO ($N = 129$)	Vehicle ($N = 51$)
Inflammatory Lesion Counts				
Mean Change	-16.6	-12.4	-11.7	-11.6
Mean Percent Change	-54.1	-44.4	-39.2	-38.2
p-value†	-	0.0044	< 0.001	.0072
Non-inflammatory Lesion Counts				
Mean Change	-24.8	-16.8	-13.8	-15.1
Mean Percent Change	-48.0	-33.9	-32.7	-34.0
p-value†	-	0.0037	<.001	0.0021
Total Lesion Counts				
Mean Change	-41.4	-29.2	-25.4	-26.7
Mean Percent Change	-50.5	-38.3	-35	-36.2
p-value†	-	< .001	< .001	.0005

† p-value is based on the main effects model with change as the response and terms for treatment, analysis center, and baseline count.

Source: Reviewer's Analysis

3.1.3.7.1 Investigator Global Assessment The definition of treatment success using the IGA scale is 'Clear' or 'Almost Clear' for both of the sensitivity analyses on the IGA.

Table 12 provides the percent of treatment successes for each treatment as well as the corresponding p-value from the CMH test stratified by analysis center. Efficacy results when imputing all the missing as failures is nearly identical to the LOCF analysis as only one subject treated with benzoyl peroxide had week 12 missing data but reached an IGA score of either 'Clear' or Almost Clear' prior to week 12. Subjects randomized to adapalene had the highest proportion of discontinuations and imputing these as successes resulted in p-value greater than 0.05.

3.1.3.7.2 Change in Lesion Counts In the sensitivity analysis on lesion counts, the response variable is the change from baseline (not ranked) with main effect terms only (treatment group, analysis center, and baseline count). In the tabular summaries, Sensitivity Analysis I refers to the case when missing data is imputed to the median change in lesion count from the IGA 'Failures'. Sensitivity Analysis II refers to when missing data is imputed as the median change in lesion count from the IGA 'Success'.

Table 13 contains the mean change, mean percent change, and p-values for each of the sensitivity analyses on lesion counts. Efficacy results using these alternate imputation strategies are consistent with the efficacy results when imputing using LOCF. All comparisons for both

Table 12: Investigator Global Results (ITT-Sensitivity): 18094

Drop Out/N (%)	EPIDUO 10/149 (6.7%)	Adapalene 17/148 (11.4%)	BPO 10/149 (6.7%)	Vehicle 7/71 (9.9%)
LOCF				
Success (%)	41 (27.5)	23 (15.5)	23 (15.4)	7 (9.9)
p-value	-	0.0079	0.0034	0.0015
Impute All Failures				
Success (%)	41 (27.5)	23 (15.5)	22 (14.8)	7 (9.9)
p-value	-	0.0079	0.0019	0.0015
Impute All Success				
Success (%)	51 (34.2)	40 (27.0)	33 (22.1)	14 (19.7)
p-value	-	0.1646	0.009	0.0196

† Source: Table 17 in Module 2.7.3 of the NDA; results reproduced by reviewer.

lesion counts and both sensitivity analyses reach statistical significance at the $\alpha = 0.05$ level.

3.1.3.8 Secondary Endpoint Results The only secondary endpoint intended for labeling claims is the percent reduction in lesion counts. As agreed upon with the Division, no multiplicity adjustment for this endpoint was required as it would be used descriptively. Point estimates for the percent reduction in lesion counts are provided in Tables 7, 8, and 9 for Study 18094.

3.1.3.9 Summary of Efficacy Findings The Division provided detailed comments about the statistical analysis of Study 18094 as the sponsor formally powered the study for efficacy. These comments were not incorporated into a revised Phase 2 protocol. However, the data was analyzed above using agreed upon statistical methods which were made in the review of the protocol for Study 18087. In the assessment of the IGA dichotomized to success using multiple definitions, EPIDUO was consistently statistically superior to each of its monads and vehicle. In the analysis of the lesion counts, the treatment effects increased when the baseline lesion count was larger. At baseline it was detected that subjects randomized to EPIDUO had the lowest number of baseline lesions. Thus, as treatment effects were smaller for lower baseline lesion counts such an imbalance in the randomization did not provide a favorable condition to demonstrate efficacy of EPIDUO. Using a main effects model with terms for treatment, baseline lesion count, and site all contrasts comparing EPIDUO to its monads and vehicle were statistically significant at the $\alpha = 0.05$ level.

Table 13: Change in Lesion Counts (ITT-Sensitivity): 18094

	EPIDUO ($N = 149$)	Adapalene ($N = 148$)	BPO ($N = 149$)	Vehicle ($N = 71$)
Inflammatory Lesion Counts				
Sensitivity Analysis I				
Mean Change	-16.8	-12.2	-10.9	-10.4
Mean Percent Change	-55.5	-42.9	-36.8	-34.2
p-value	-	$< .001$	$< .001$	$< .001$
Sensitivity Analysis II				
Mean Change	-17.2	-13.2	-11.5	-10.9
Mean Percent Change	-57	-45.9	-38.5	-35.9
p-value	-	$< .001$	0.0031	$< .001$
Non-Inflammatory Lesion Counts				
Sensitivity Analysis I				
Mean Change	-24.2	-16.5	-13.8	-14.6
Mean Percent Change	-48.4	-32.9	-32.3	-31
p-value	-	$< .001$	$< .001$	$< .001$
Sensitivity Analysis II				
Mean Change	-24.9	-17.9	-14.9	-15.7
Mean Percent Change	-50.1	-36.1	-34.9	-33.4
p-value	-	$< .001$	$< .001$	$< .001$

Sensitivity Analysis I refers to the case when missing data is imputed to the median change in lesion count from the IGA 'Failures'. Sensitivity Analysis II refers to when missing data is imputed as the median change in lesion count from the IGA 'Success'.
Source: Reviewer's Analysis.

3.1.4 Efficacy: Study 18087

Study 18087 is one of two Phase 3 trials, however Study 18088 was ongoing at the time of submission and therefore not included. In the sample size calculation for Study 18087, the sponsor relied on the efficacy results of 18094. Specifically the percent success on the IGA was the most influential endpoint for deriving sample size. The sponsor assumed a drop out rate of 15%, and response rates of 25% for EPIDUO and 15% for each of the monads. With 414 subjects per treatment arm the study would have greater than 90% power when testing at the two-sided $\alpha = 0.05$ level.

3.1.4.1 Study Design Study 18087 was a multi-center, randomized, double-blind, parallel, active- and vehicle-controlled study to evaluate the safety and efficacy of EPIDUO compared with adapalene, benzoyl peroxide, and vehicle in the treatment of acne vulgaris when applied once daily. Male and female subjects, age 12 years or older, with a baseline IGA score of 'Moderate' and 20 to 50 inflammatory lesions and 30 to 100 non-inflammatory lesions on the face were randomized at the baseline visit, instructed to apply study medication once daily in the evening to treat the affected areas for up to 12 weeks. 1668 subjects were recruited and randomized in 1:1:1:1 ratio to EPIDUO, adapalene, benzoyl peroxide, and vehicle, respectively at 60 centers in the U.S, Puerto Rico, and Canada.

The treatment period consisted of 12 weeks with 6 study visits occurring at baseline, week 1, week 2, week 4, week 8, and week 12. The primary timepoint for efficacy evaluation is at the week 12 visit which was planned to occur within a 7 day treatment window. The protocol stated efficacy objective of the trial was to demonstrate the clinical efficacy of EPIDUO compared with adapalene, benzoyl peroxide, and vehicle as assessed by the primary efficacy criteria of IGA success ("Clear" or "Almost Clear") and change in non-inflammatory and inflammatory lesion counts. In addition, the trial was designed to assess the safety and tolerability profile of EPIDUO as assessed by the local tolerability parameters (erythema, scaling, dryness and stinging/burning) and adverse event reporting.

3.1.4.2 Endpoints As agreed upon in the review of the protocol for Study 18087 the following are considered to be the co-primary endpoints to establish the efficacy for the indication of acne vulgaris.

- Success rate, the percentage of subjects with "0 = Clear" or "1 = Almost Clear" on the Investigators Global Assessment (0 to 4 scale) at week 12
- Changes in inflammatory and non-inflammatory lesion counts from baseline to week 12

Percent changes of the lesion counts are the only secondary endpoints intended for labeling claims. In the review of SN054 (stamp date: 04/20/2007) the Division agreed that these sec-

ondary endpoints could be included in the label if the primary endpoint for change in lesion counts meets statistical criteria without multiplicity adjustment.

3.1.4.3 Patient Disposition and Baseline Characteristics

3.1.4.3.1 Patient Disposition A total of 239 out of 1668 subjects (14.3%) discontinued from Study 18087; reason recorded for subject discontinuation is provided in Table 14. The most prevalent reason for subject discontinuation was due to "lost to follow-up" which accounted for 116 subjects (7.0%) who discontinued. 11 of the 22 subjects that discontinued treatment due to an adverse event were treated with EPIDUO.

Table 14: Subject Disposition (Study 18087)

	EPIDUO ($N = 415$)	Adapalene ($N = 420$)	Benzoyl Peroxide ($N = 415$)	Vehicle ($N = 418$)
Complete Trial	347 (83.6)	363 (86.4)	372 (89.6)	347 (83.0)
Discontinued	68 (16.4)	57 (13.6)	43 (10.4)	71 (17.0)
Lack of Efficacy	1 (0.2)	2 (0.5)	0 (0.0)	1 (0.2)
Adverse Event	11 (2.7)	4 (1.0)	5 (1.2)	2 (0.5)
Subject Request	21 (5.1)	17 (4.0)	18 (4.3)	30 (7.2)
Protocol Violation	0 (0.0)	1 (0.2)	0 (0.0)	1 (0.2)
Lost to Follow-up	31 (7.5)	32 (7.6)	19 (4.6)	34 (8.1)
Other	1 (0.2)	0 (0.0)	0 (0.0)	1 (0.2)
Pregnancy	3 (0.7)	1 (0.2)	1 (0.2)	2 (0.5)

Source: Sponsor's Study Report Table 11; results reproduced by reviewer.

3.1.4.3.2 Baseline Demographic Factors The baseline demographics for age, gender, and race are provided in Table 32 in the Appendix. Overall, the mean age of subjects was 18 years old; approximately two-thirds of subjects were identified as Caucasian; and 51% of subjects were female. There was no imbalance of any of the demographic factors between treatment arms.

3.1.4.3.3 Baseline Prognostic Factors This exploratory analysis examines the same baseline characteristics of acne vulgaris as performed for Study 18094. Analysis results are provided in Table 15. Almost all subjects had baseline IGA scores of 'Moderate' which is expected as the inclusion criteria required an IGA score of 'Moderate'. In terms of the baseline lesion counts, subjects randomized to EPIDUO had lower median counts for both inflammatory and non-inflammatory lesion counts than the other treatment arms. Looking at an empirical cumulative distribution of the lesion counts by treatment group it shows that the proportion of

subjects with any given lesion count is always smallest for subjects randomized to EPIDUO (empirical cumulative distribution plot provided in the Appendix, Figure 20). As in Study 18094 the baseline distribution of EPIDUO is similar in shape with a shift to the left.

Table 15: Baseline Prognostic Factors by Treatment (Study 18087)

	N	EPIDUO $N = 415$	Adapalene $N = 420$	Benzoyl Peroxide $N = 415$	Vehicle $N = 418$
IGA : Moderate	1666†	100% (415)	100% (420)	100% (414)	100% (416)
Inflammatory	1668	5 10 20	8 14 23	7 12 21	10 19 28
Non-inflammatory	1668	11.00 22.00 38.00	14.00 25.00 41.00	15.00 27.00 45.00	21.00 32.50 49.75
Total	1668	18.00 34.00 57.00	24.00 41.00 64.25	24.00 42.00 68.00	35.25 53.00 76.00
Burning : None	1666‡	97% (402)	97% (406)	97% (403)	96% (401)
Mild		2% (10)	3% (13)	3% (12)	3% (14)
Moderate		0% (2)	0% (1)	0% (0)	0% (2)
Severe		0% (0)	0% (0)	0% (0)	0% (0)
Dryness : None	1666‡	88% (364)	87% (366)	87% (359)	86% (359)
Mild		12% (49)	12% (51)	12% (51)	13% (55)
Moderate		0% (1)	1% (3)	1% (5)	1% (3)
Severe		0% (0)	0% (0)	0% (0)	0% (0)
Erythema : None	1666‡	78% (322)	78% (328)	73% (305)	77% (322)
Mild		16% (66)	17% (73)	20% (82)	19% (80)
Moderate		6% (24)	5% (19)	7% (28)	3% (13)
Severe		0% (2)	0% (0)	0% (0)	0% (2)
Scaling : None	1666‡	92% (382)	91% (383)	91% (379)	92% (383)
Mild		7% (30)	9% (37)	8% (34)	8% (33)
Moderate		0% (2)	0% (0)	0% (2)	0% (1)
Severe		0% (0)	0% (0)	0% (0)	0% (0)

a b c represent the lower quartile *a*, the median *b*, and the upper quartile *c* for continuous variables. *N* is the number of non–missing values. Numbers after percents are frequencies.

† Note that two subjects did not have baseline IGA scores (ID's: 92007 and 92381)

‡ Note that two subjects did not have baseline scores for local skin reactions(ID's: 91585 and 92381)

Source: Reviewer's analysis.

The local skin reactions: burning, dryness, erythema, and scaling were ranked from none (0) to severe (5). At baseline a large majority of the subjects scored dryness, burning, and scaling as 'None'. Around 15% of subjects scored erythema as 'Mild' and even a small percentage had scores of 'Moderate'. Overall the distributions were balanced across treatment arms for local skin reactions.

3.1.4.4 Populations Analyzed The Intent-to-Treat population (ITT) includes all subjects randomized and dispensed study medication. Subjects with major protocol deviations were excluded from the Per-Protocol (PP) population. Table 16 depicts the number of subjects included in the two analysis populations for each of the studies. Of note is that subject 92381 enrolled in Study 18087 had no baseline or post-baseline IGA recorded, though lesion counts were assessed. Thus, this subject is not included in the ITT population for IGA analysis but is included in the ITT population for the analysis of change in lesion counts.

Table 16: Summary of Data Sets Analyzed

	EPIDUO	Adapalene	BPO	Vehicle
ITT Population	415	420	415	418
PP Population	319	347	346	323

Source: Sponsor's Study Report Table 5; results reproduced by reviewer.

3.1.4.5 Primary Endpoint Results (ITT-LOCF) The section on efficacy is broken down into two sections. The first examines efficacy results according to the investigator global assessment whereas the second examines efficacy results based upon change in lesion counts.

3.1.4.5.1 Investigator Global Assessment In the analysis of percent success on IGA, EPIDUO was superior to each monad and its vehicle in Study 18087. A success is defined for subjects that receive an IGA score of 0 (clear) or 1 (almost clear) at week 12. To test the superiority of EPIDUO to the other three treatment arms, a Cochran-Mantel-Haenszel (CMH) test was carried out with adjustments for analysis center. The results of the CMH test are provided in Table 17. Based on the CMH tests, EPIDUO is superior to each monad and vehicle in Study 18087 at the statistical significance level of $\alpha = 0.05$.

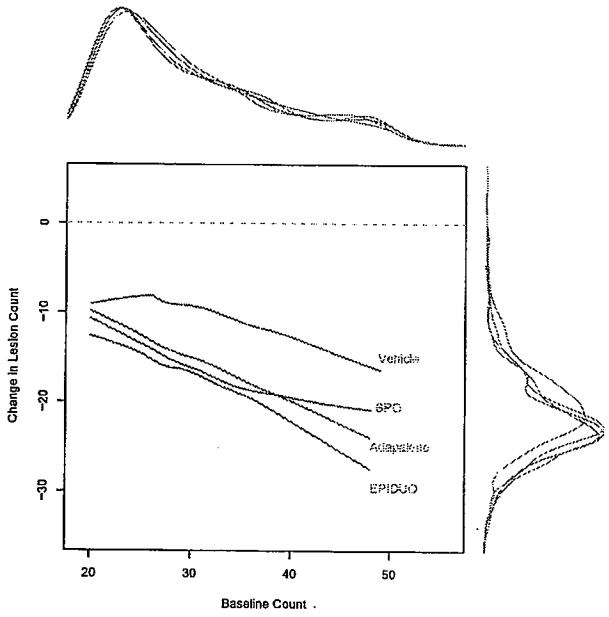
Table 17: Investigator Global Results (ITT-LOCF)

	EPIDUO ($N = 415$)	Adapalene ($N = 420$)	BPO ($N = 415$)	Vehicle ($N = 417$)
Success (%)	125 (30.1)	83 (19.8)	92 (22.2)	47 (11.3)
p-value	-	<.001	0.0062	<.001

Source: Sponsor's Study Report Tables 10 and 11; results reproduced by reviewer.

3.1.4.5.2 Change in Lesion Counts Figure 5 depicts loess fits for the inflammatory lesions at baseline and the change in inflammatory lesion count along with univariate density estimates of the distributions for each treatment group. The loess fits show a trend of a decrease in the inflammatory lesion counts for all treatment groups (i.e. all loess fits are below the line depicting no change). In general, EPIDUO had the lowest inflammatory lesion counts at the end of treatment regardless of the baseline inflammatory lesion count. The slopes of the smoothed regression lines is roughly parallel for EPIDUO, adapalene, and vehicle. However, the slope of the line for benzoyl peroxide tends to level off for baseline inflammatory lesion counts greater than 35.

Figure 5: Inflammatory Lesion Counts (Study 18087)



In the modeling of change in inflammatory lesion counts, neither of the interaction terms were significant ($\alpha = 0.05$). Thus, as defined in the protocol a main effects model was fit with terms for treatment, analysis site, and baseline count. The main effects model as shown in Equation 1 was used to test for significant treatment effects using contrasts to compare EPIDUO to benzoyl peroxide, adapalene, and vehicle.

$$\Delta \text{ INF Lesion Ct.} = \text{Treatment} + \text{Analysis Center} + \text{Baseline INF Lesion Ct.} \quad (1)$$

Efficacy results for inflammatory lesion counts are listed in Table 18. Results are provided for both the ranked and unranked data. The p-value on the ranked data is above the $\alpha = 0.05$

level for the comparison of EPIDUO to benzoyl peroxide though the p-value is 0.0387 for the unranked data.

Table 18: Change in Inflammatory Lesion Counts (ITT-LOCF): 18087

	EPIDUO ($N = 415$)	Adapalene ($N = 420$)	BPO ($N = 415$)	Vehicle ($N = 418$)
Inflammatory Lesion Counts				
Mean Change	-15.4	-12.3	-13.7	-8.7
Mean Percent Change	-53.4	-41.7	-47.6	-30.2
p-value† (ranked)	-	< 0.001	0.068	< 0.001
p-value‡ (untransformed)	-	< 0.001	0.0387	< 0.001

† Sponsor's analysis using a main effects model with the ranked data; results reproduced by the reviewer.

‡ Reviewer's analysis using a main effects model on the unranked data.

Figure 6 depicts a smoothed regression line for the non-inflammatory lesions at baseline and change in non-inflammatory lesion count along with univariate density estimates of the distributions for each treatment group. The smoothed regression lines show a trend of a decrease in the non-inflammatory lesion counts for each treatment group (i.e. all smoothed regression lines are below the line depicting no change). The smoothed regression lines show that EPIDUO did not consistently have a greater estimated decrease in the number of non-inflammatory lesions than adapalene throughout the whole range of baseline non-inflammatory lesion counts.

As seen in Figure 6 a significant baseline by treatment interaction was found in the modeling of the non-inflammatory lesions ($p < 0.001$). The sponsor's analysis relies on a main effects model with the ranked data[4]. However, as the Shapiro-Wilk's test is not significant for the sponsor's model, the review also provides efficacy results on the unranked data using a model with the interaction of treatment and baseline non-inflammatory lesion count.

The predicted change in non-inflammatory lesion counts for a given baseline non-inflammatory lesion count is depicted in Figure 7 for each treatment in analysis center 42. The figure suggests that the predicted change in non-inflammatory lesions for vehicle is the cause of the significant interaction term. Overall, the predicted change in non-inflammatory lesions is greatest for subjects randomized to EPIDUO.

[4]Note that when using ranked data, the interaction term was no longer significant.

Figure 6: Non-Inflammatory Lesion Counts (Study 18087)

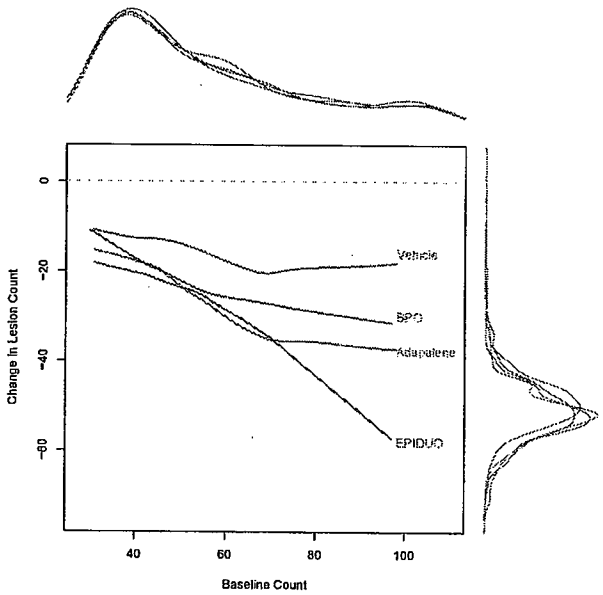
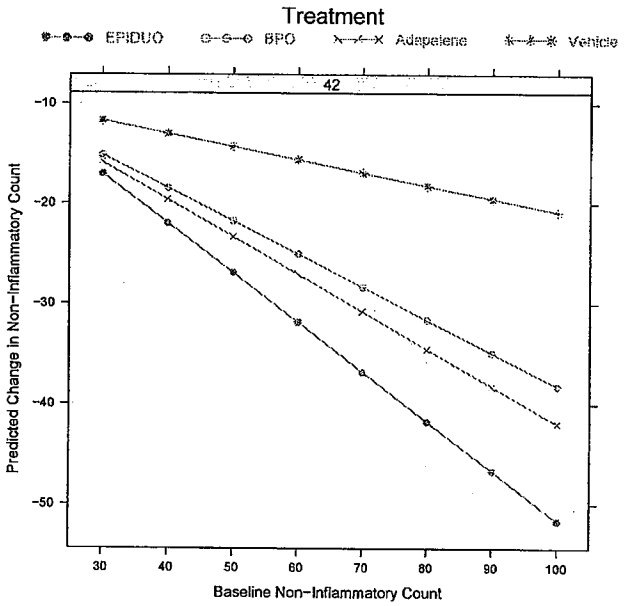


Figure 7: Non-Inflammatory Lesion Counts (Study 18087)



A summary of the change as well as the percent reduction in non-inflammatory lesion counts is provided in Table 19. The table also provides p-values for testing the main treatment effect for a model with the main effects only not the model as described above with the interaction term. Note that the use of the ranked data results in a larger p-value than the model on the unranked data (results not shown).

Table 19: Change in Non-Inflammatory Lesion Counts (ITT-LOCF): 18087

	EPIDUO ($N = 415$)	Adapalene ($N = 420$)	BPO ($N = 415$)	Vehicle ($N = 418$)
Mean Change	-24.6	-21.0	-19.2	-11.3
Mean Percent Change	-48.1	-40.8	-37.2	-23.2
p-value[†](ranked)	-	0.048	< 0.001	< 0.001
p-value[‡](untransformed)	-	0.0009	0.0002	< 0.001

[†] Sponsor's analysis using a main effects model with the ranked data; results reproduced by the reviewer.

[‡] Reviewer's analysis using a main effects model with the the unranked data.

A summary of the change as well as the percent reduction in total lesion counts is provided in Table 20. The table also provides p-values for testing the main treatment effect for a model with the main effects only. In terms of change in number of total lesions, EPIDUO was statistically superior to each of its monads and vehicle.

Table 20: Change in Total Lesion Counts (ITT-LOCF): 18087

	EPIDUO ($N = 415$)	Adapalene ($N = 420$)	BPO ($N = 415$)	Vehicle ($N = 418$)
Mean Change	-39.9	-33.3	-33.0	-20.0
Mean Percent Change	-50.	-41.3	-41.2	-26.1
p-value	-	0.0003	0.0004	< 0.001

Source: Reviewer's analysis using an ANCOVA model with main effects only on the unranked data.

3.1.4.6 Primary Endpoint Results (PP-LOCF)

3.1.4.6.1 Investigator Global Assessment Based upon the definition of IGA success as 'Clear' or 'Almost clear', the efficacy results on the PP population are listed in Table 21. Efficacy results in the PP population are consistent with the ITT population.

Table 21: Investigator Global Results (PP-LOCF): 18087

	EPIDUO ($N = 319$)	Adapalene ($N = 347$)	BPO ($N = 346$)	Vehicle ($N = 323$)
Success (%)	112 (35.1)	74 (21.3)	81 (23.4)	43 (13.3)
p-value[†]	-	<.001	<.001	<.001

[†] Source: Sponsor's Study Report Table EFF 5.2; results reproduced by reviewer.

3.1.4.6.2 Change in Lesion Counts For the assessment of absolute change in lesion count in the PP population, the main effects model is fit with absolute change in lesion counts as the response and terms for treatment, analysis center, and baseline lesion count (see equation 1). Note that the sponsor reported efficacy results using ranks on the lesion counts, but as previously discussed this does not improve the normality of the residuals. As such efficacy results provided below are based upon data that is not ranked.

Efficacy results for the change in inflammatory, non-inflammatory, and total lesion counts is provided in Table 22 for the PP population. As seen in Study 18094, the observed point estimates for change and percent change in Study 18087 show more reduction in the baseline lesions in the PP population than the ITT population. Of note is that the treatment effect comparing EPIDUO to benzoyl peroxide for inflammatory lesions in the PP population is 2.6 whereas it was 1.7 in the ITT population.

3.1.4.7 Missing Data Sensitivity Analysis As a sensitivity analysis to the primary method of data imputation as LOCF, two alternate imputation approaches were provided in the protocol and agreed upon with the Division. For the IGA the alternate imputation approached are: (a) impute all missing week 12 data as failures and (b) impute all missing week 12 data as successes. For the change in lesion counts the alternate imputation approaches are: (a) impute the median change in lesion count from the IGA 'Failures' for each treatment group and (b) impute the median change in lesion count from the IGA 'Success' for each treatment group.

Table 22: Change in Lesion Counts (PP-LOCF): 18087

	EPIDUO ($N = 319$)	Adapalene ($N = 347$)	BPO ($N = 346$)	Vehicle ($N = 323$)
Inflammatory Lesion Counts				
Mean Change	-17.2	-13.0	-14.6	-9.8
Mean Percent Change	-59.6	-44.3	-50.2	-33.3
p-value†	-	< .001	0.0018	< .001
Non-inflammatory Lesion Counts				
Mean Change	-27.2	-22.8	-20.8	-11.3
Mean Percent Change	-52.9	-44.4	-40.0	-23.3
p-value†	-	0.0054	< .001	< .001
Total Lesion Counts				
Mean Change	-44.5	-35.8	-35.4	-21.1
Mean Percent Change	-55.3	-44.5	-44	-27.5
p-value†	-	< .001	< .001	< .001

† p-value is based on the main effects model with change as the response and terms for treatment, analysis center, and baseline count.

Source: Reviewer's Analysis

3.1.4.7.1 Investigator Global Assessment The definition of treatment success using the IGA scale is 'Clear' or 'Almost Clear' for both of the sensitivity analyses on the IGA. The proportion of discontinuations is similar for all treatment arms in Study 18087 as in Study 18094. Consequently, the p-values for all the proposed imputation strategies are consistent between the two studies which show EPIDUO is superior to each of its monads and vehicle. Analysis results are shown in Table 23.

3.1.4.7.2 Change in Lesion Counts In the sensitivity analysis on lesion counts, the response variable is the change from baseline (not ranked) with main effect terms only (treatment group, analysis center, and baseline count). In the tabular summaries, Sensitivity Analysis I refers to the case when missing data is imputed to the median change in lesion count from the IGA 'Failures'. Sensitivity Analysis II refers to when missing data is imputed as the median change in lesion count from the IGA 'Success'.

Table 24 contains the mean change, mean percent change, and p-values for each of the sensitivity analyses on lesion counts. Efficacy results using these alternate imputation strategies are consistent with the efficacy results when imputing using LOCF. Overall, efficacy results based on the sensitivity analysis of the missing data are consistent with the primary analysis results.

Table 23: Investigator Global Results (ITT-Sensitivity): 18087

Drop Out/N (%)	EPIDUO 68/415 (16.4%)	Adapalene 57/420 (13.6%)	BPO 43/415 (10.4%)	Vehicle 71/418 (17.0%)
LOCF				
Success (%)	125 (30.1)	83 (19.8)	92 (22.2)	47 (11.3)
p-value	-	<.001	0.0062	<.001
Impute All Failures				
Success (%)	121 (29.2)	79 (18.8)	91 (21.9)	46 (11)
p-value	-	<.001	0.012	<.001
Impute All Success				
Success (%)	188 (45.3)	136 (32.4)	134 (32.3)	114 (27.3)
p-value	-	<.001	<.001	<.001

† Source: Table 19 in Module 2.7.3 of the NDA; results reproduced by reviewer.

3.1.4.8 Secondary Endpoint Results The only secondary endpoint intended for labeling claims is the percent reduction in lesion counts. As agreed upon with the Division, no multiplicity adjustment for this endpoint was required as it would be used descriptively. Point estimates for the percent reduction in lesion counts are provided in Tables 18, 19, and 20.

Table 24: Change in Lesion Counts (ITT-Sensitivity): 18087

	EPIDUO ($N = 415$)	Adapalene ($N = 420$)	BPO ($N = 415$)	Vehicle ($N = 417$)
Inflammatory Lesion Counts				
Sensitivity Analysis I				
Mean Change	-16.2	-14.7	-13.1	-9.3
Mean Percent Change	-56.6	-51	-44.7	-32.8
p-value	-	$< .001$	0.0293	$< .001$
Sensitivity Analysis II				
Mean Change	-17.4	-15.4	-14.2	-10.9
Mean Percent Change	-60.7	-53.8	-48.5	-38.7
p-value	-	$< .001$	0.0073	$< .001$
Non-Inflammatory Lesion Counts				
Sensitivity Analysis I				
Mean Change	-25.8	-20.6	-22.5	-11.8
Mean Percent Change	-51.1	-40.3	-44.4	-24.5
p-value	-	0.0116	$< .001$	$< .001$
Sensitivity Analysis II				
Mean Change	-27.5	-21.9	-23.9	-14.1
Mean Percent Change	-55.1	-43.3	-47.4	-29.4
p-value	-	0.0051	$< .001$	$< .001$

Sensitivity Analysis I refers to the case when missing data is imputed to the median change in lesion count from the IGA 'Failures'. Sensitivity Analysis II refers to when missing data is imputed as the median change in lesion count from the IGA 'Success'. Source: Reviewer's Analysis.

3.1.4.9 Sensitivity Analysis of Lesion Counts In the fitting of the ANCOVA models for the lesion counts in Study 18087 the distribution of the residuals were not normally distributed regardless of whether the counts are ranked or not. As sample sizes for this trial are relatively large the distributional assumption of normality may not greatly impact the parameter estimates. However, the following is a sensitivity analysis to assess a transformation on the response to improve the properties of the ANCOVA model.[5]

In this sensitivity analysis on the ITT-LOCF population, rather than use change from baseline as the response, the end of treatment lesion count is defined as the response variable which is denoted by y. In addition, main effect terms only are included in the ANCOVA model. In this sensitivity analysis the response y is transformed in such a way to improve the properties of the ANCOVA model. Similar to using the Box-Cox transformation, the one parameter family of transformations considered is the following.

$$f(y, \alpha) = log(y + \alpha) \tag{2}$$

The profile log likelihood for α is calculated using the following equation.

$$\hat{L}(\alpha) = \text{const} - \frac{n}{2} log\text{RSS}\{log(y + \alpha)\} - \sum log(y + \alpha) \tag{3}$$

Here, RSS refers to the residual sum of squares for the regression of $log(y + \alpha)$. Using multiple choices of α the likelihood is calculated and the α with maximum value of $\hat{L}(\alpha)$ is selected.

3.1.4.9.1 Inflammatory Lesion Counts Figure 8 depicts several diagnostics of the model without transforming the response variable, end of treatment inflammatory lesion count. In this figure several outliers are detected though these are relatively small in number. The normal Q-Q plot shows that the right tail of the distribution is higher than expected. Using the Shapiro-Wilk normality test on the residuals of this model resulted in a p-value < 0.0001. In addition, the scatterplot of the fitted values versus the standardized residuals suggests a potential problem with nonconstant variance.

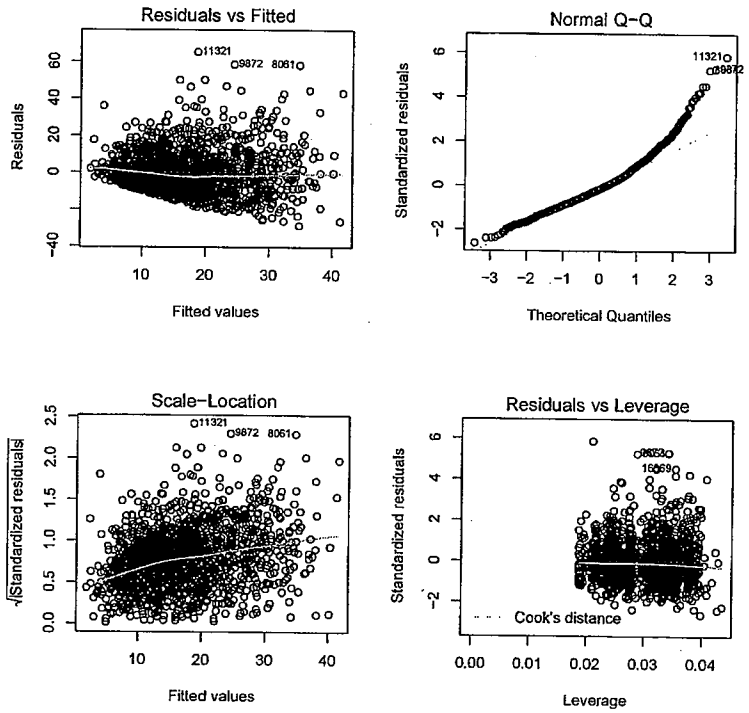
Using the methods above, the estimated value of α was five[6]. Thus the transformed response was $log(y+5)$, where y =number of inflammatory lesions at end of treatment. This response was then fit to the main effects model. The diagnostics of this model with the transformed response are shown in Figure 9. Based upon the diagnostic plots, the model using the transformed response shows no apparent violations of model assumptions improving upon the model with the untransformed response. Using the Shapiro-Wilk normality test on the residuals of this model resulted in a p-value of 0.0044 which is likely due to the lack of power for such a test.

The estimated regression coefficients and the corresponding test statistics are provided in Table 25 for both the untransformed and transformed response. In all cases the test statistic

[5]Diagnostic plots of the model based on ranks is provided in Appendix Section A.2.

[6]Note that using the same procedure on the inflammatory lesion count data from Study 18094, α was also 5.

Figure 8: Model Diagnostics for Untransformed Inflammatory Lesion Count



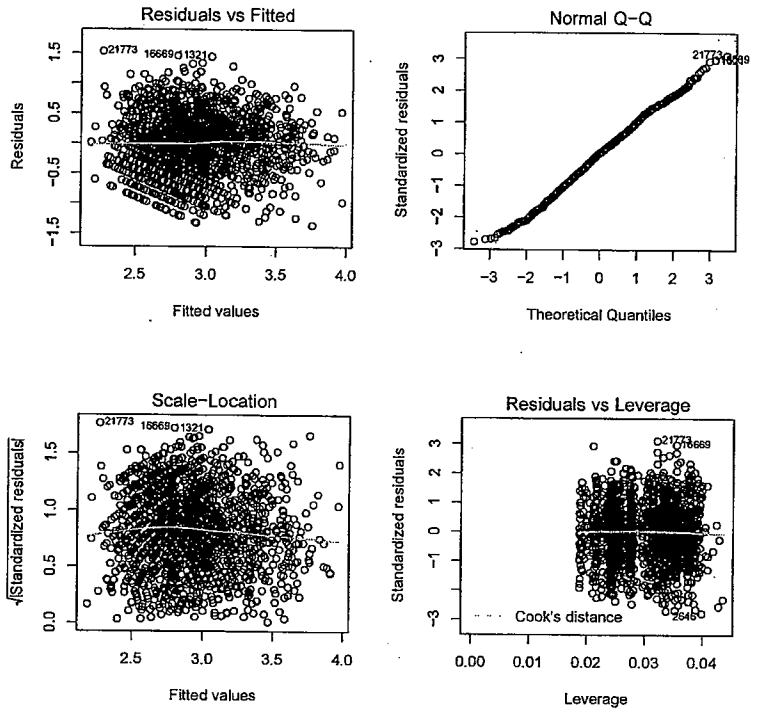
increased for the transformed response though it was not by a large margin. In both analysis, the comparison of EPIDUO to each of its monads and vehicle all resulted in p-values less than 0.05.

Table 25: Regression Parameter Estimates: Inflammatory Lesion Counts

	Estimate	Std. Error	T-Value	p-value
Untransformed Model				
EPIDUO vs. BPO	1.62	.7834	2.07	.0387
EPIDUO vs. Adapalene	3.15	.7812	4.03	< 0.001
EPIDUO vs. Vehicle	6.67	.7821	8.53	< 0.001
Transformed Model				
EPIDUO vs. BPO	0.08	.0347	2.33	.0199
EPIDUO vs. Adapalene	0.17	.0346	4.90	< 0.001
EPIDUO vs. Vehicle	0.33	.0346	9.40	< 0.001

Source: Reviewer's analysis.

Figure 9: Model Diagnostics for Transformed Inflammatory Lesion Count

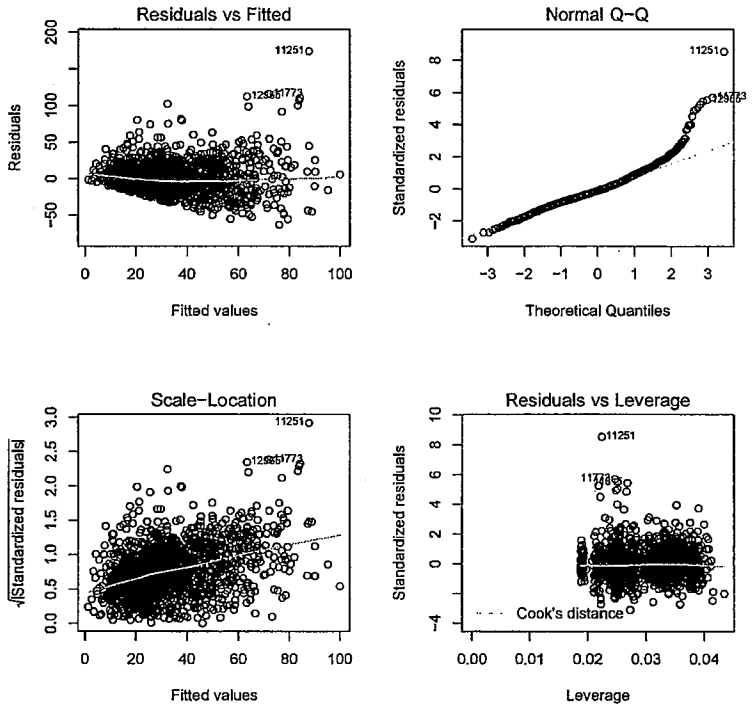


3.1.4.9.2 Non-Inflammatory Lesion Counts Figure 10 depicts several diagnostics of the model without transforming the response variable, end of treatment non-inflammatory lesion count. In this figure several outliers are detected though these are relatively small in number. The normal Q-Q plot shows that the right tail of the distribution is higher than expected. Using the Shapiro-Wilk normality test on the residuals of this model resulted in a p-value < 0.0001. In addition, the scatterplot of the fitted values versus the standardized residuals suggests a potential problem with nonconstant variance.

Using the methods above, the estimated value of α was eleven[7]. Thus the transformed response was $log(y + 11)$, where y =number of non-inflammatory lesions at end of treatment. This response was then fit to the main effects model. The diagnostics of this model with the transformed response are shown in Figure 11. Based upon the diagnostic plots, the model using the transformed response shows no apparent violations of model assumptions improving upon the model with the untransformed response. Using the Shapiro-Wilk normality test on the residuals of this model resulted in a p-value of 0.0435 which is likely due to the lack of power

[7]Note that using the same procedure on the non-inflammatory lesion count data from Study 18094, α was also near 11.

Figure 10: Model Diagnostics for Untransformed Non-Inflammatory Lesion Count



for such a test.

The estimated regression coefficients and the corresponding test statistics are provided in Table 26 for both the untransformed and transformed response. In all cases the test statistic increased for the transformed response though it was not by a large margin. In both analysis, the comparison of EPIDUO to each of its monads and vehicle all resulted in p-values less than 0.05.

The sensitivity analysis demonstrated that using the transform, $log(y + \alpha)$ for both the inflammatory and non-inflammatory lesion counts resulted in regression models with no apparent violations of model assumptions. In addition, the estimate of α was similar between Study 18087 and Study 18094 (results not shown for Study 18094). Efficacy results from this sensitivity analysis were consistent with those of the untransformed data using a main effects model.

Figure 11: Model Diagnostics for Transformed Non-Inflammatory Lesion Count

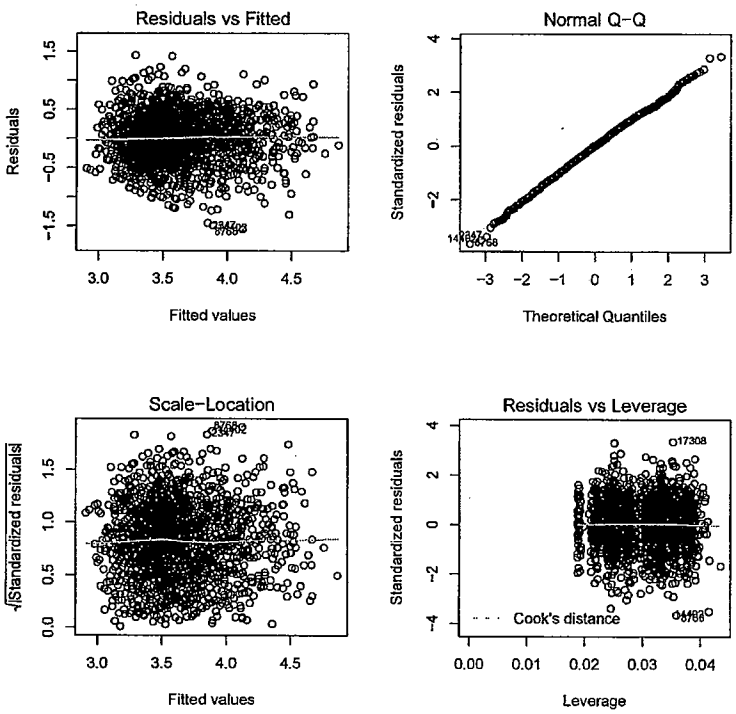


Table 26: Regression Parameter Estimates: Non-Inflammatory Lesion Counts

	Estimate	Std. Error	T-Value	p-value
Untransformed Model				
EPIDUO vs. BPO	5.43	1.4326	3.79	.0002
EPIDUO vs. Adapalene	3.74	1.4286	2.62	.0088
EPIDUO vs. Vehicle	13.14	1.4301	9.19	< 0.001
Transformed Model				
EPIDUO vs. BPO	0.14	.0302	4.62	< 0.001
EPIDUO vs. Adapalene	0.09	.0301	3.11	0.0019
EPIDUO vs. Vehicle	0.30	.0302	9.82	< 0.001

Source: Reviewer's analysis.

3.1.4.10 Summary of Efficacy Findings Study 18087 was powered at over 90% using estimates of treatment effects from Study 18094. The Division and sponsor reached agreements on the statistical analysis plan. Based upon the pre-specified criteria, EPIDUO was superior to each of its monads and vehicle for the co-primary IGA endpoint as well as for the change in

non-inflammatory lesion endpoint. For the co-primary endpoint, change in inflammatory lesion count, the comparison of EPIDUO to benzoyl peroxide did not reach statistical significance ($p = 0.068$) for the protocol defined primary analysis method. This analysis was based on ranking the data prior to fitting the ANCOVA model as ranks were used due to the significance of Shapiro-Wilk's normality test. An examination of the ANCOVA model diagnostics did not reveal an added benefit to using the ranks over the untransformed data in which case the analysis based on the untransformed data reached statistical significance ($p = 0.0387$). In addition, a sensitivity analysis using a transformation of the response, end of treatment inflammatory lesion count, with model terms for treatment, site, and baseline count showed more desirable properties in terms of assessing the model diagnostics. This sensitivity analysis also showed a statistically significant difference in inflammatory lesions between EPIDUO and benzoyl peroxide ($p = 0.0199$). Thus, the collective evidence suggests that EPIDUO is statistically superior to each of its monads and vehicle for *all* co-primary endpoints.

3.1.5 Efficacy: Study 18089

While Study 18089 is open-label and does not form the basis for an efficacy claim, the following section provides a description of the study as well as a summary of the trial efficacy results. It should be noted that such a trial is mainly used to assess the long-term safety of EPIDUO.

3.1.5.1 Study Design Study 18089 was a multi-center, long-term, open-label study with the primary objective of assessing the safety of EPIDUO and secondary objective of assessing the efficacy in the treatment of acne vulgaris when applied once daily for up to 12 *months.* Male and female subjects, age 12 years or older, with 20 to 50 inflammatory lesions and 30 to 100 non-inflammatory lesions on the face were eligible to be enrolled. Subjects who enrolled were instructed to apply study medication once daily in the evening to treat the affected areas. Subjects who achieved complete clearance of inflammatory and non-inflammatory lesions during the study were to stop treatment. Subjects with any recurrence of lesions after stopping treatment were allowed to resume treatment as directed by the investigator. The study enrolled 452 subjects from 28 centers in the U.S.

The treatment period consisted of up 12 months with 10 study visits occurring at baseline, week 1, week 2, month 1, month 2, month 4, month 6, month 8, month 10, and month 12. While the primary objective was to assess the safety of EPIDUO, the secondary objective was to assess the efficacy of EPIDUO with long-term use. Note the trial was designed to actively assess the safety and tolerability profile of EPIDUO as measured by the local tolerability parameters (erythema, scaling, dryness and stinging/burning).

3.1.5.2 Endpoints For efficacy assessment, no investigator global assessment was made during the study. Rather, the efficacy endpoints consisted of the inflammatory and non-inflammatory lesion counts. For the assessment of local skin reactions, the four parameters: erythema, scaling, dryness, stinging/durning, were each recorded on a four point scale (0 = None, 1 = Mild, 2 = Moderate, 3 = Severe).

3.1.5.3 Patient Disposition and Baseline Characteristics

3.1.5.3.1 Patient Disposition A total of 125 out of 452 subjects (27.7%) discontinued from Study 18089; reason recorded for subject discontinuation is provided in Table 27. The most prevalent reason for subject discontinuation was due to "subject request" which accounted for 70 subjects (15.5%) who discontinued. No subjects were recorded as withdrawing due to a lack of efficacy though Subject 255 who was listed as withdrawing due to an adverse event actually had worsening of acne as the reason for discontinuation according to the CRF.

Table 27: Subject Disposition (Study 18089)

	EPIDUO ($N = 452$)
Completed Trial	327 (72.3)
Discontinued	125 (27.7)
Adverse Event	9 (2.0)
Subject Request	70 (15.5)
Protocol Violation	2 (0.4)
Lost to Follow-up	42 (9.3)
Pregnancy	2 (0.4)

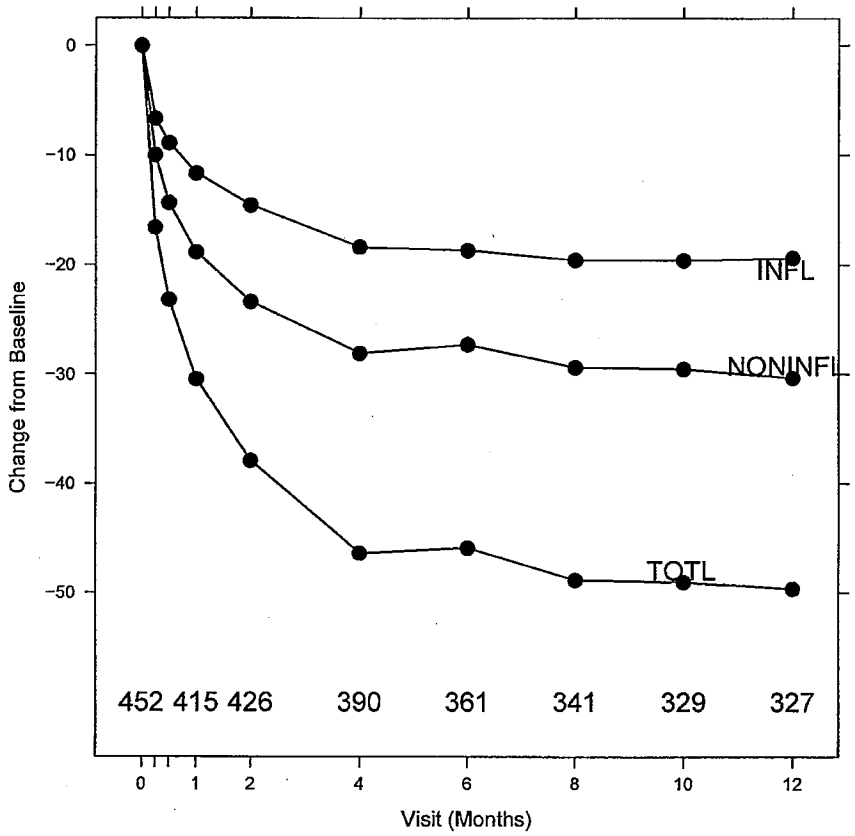
Source: Study Report Table 7; results reproduced by reviewer.

3.1.5.3.2 Baseline Demographic Factors The baseline demographics for age, gender, and race are provided in Table 33 in the Appendix. Overall, the mean age of subjects was 18 years old; approximately three-quarter of subjects were identified as Caucasian; and 51% of subjects were female.

3.1.5.3.3 Efficacy Highlights As Study 18089 is open-label with no comparator arm. Figure 12 is a graphical depiction of the mean change in lesion counts over time. Note that the calculations are based on the observed cases where the number of subjects present at the visit is shown at the bottom of the graphic. From the figure it can be seen that after 4 months of

treatment there is little additional improvement as the curves tend to become horizontal after month 4.

Figure 12: Change in Lesion Counts Across Time (18089)



3.2 Evaluation of Safety

The review of of short-term safety is based on Study 18094 and Study 18087. Assessment of long-term safety is based on Study 18089. The adverse events are coded using the MedDRA dictionary version 6.1. For subjects that experienced an AE multiple times, only a single instance is used in the tabulations below.

3.2.1 Short-term Safety Evaluation

The safety data collected from Studies 18094 and 18087 were combined to assess the safety of short-term use of the drug products. This resulted in 564, 568, 564, and 489 subjects being

exposed to EPIDUO, adapalene, benzoyl peroxide, and vehicle, respectively.

3.2.1.1 MedDRA Tabulation Table 28 contains the MedDRA preferred terms which were reported in at least 3.0% of subjects grouped according to the system organ classification (SOC). Comparing EPIDUO to its monads and vehicle, there appears to be a higher incidence of application site burning and dry skin associated with EPIDUO. Note that these two preferred terms are also likely closely related to the active assessment of the local skin reactions.

Table 28: Adverse Events by System Organ Class and Preferred Term

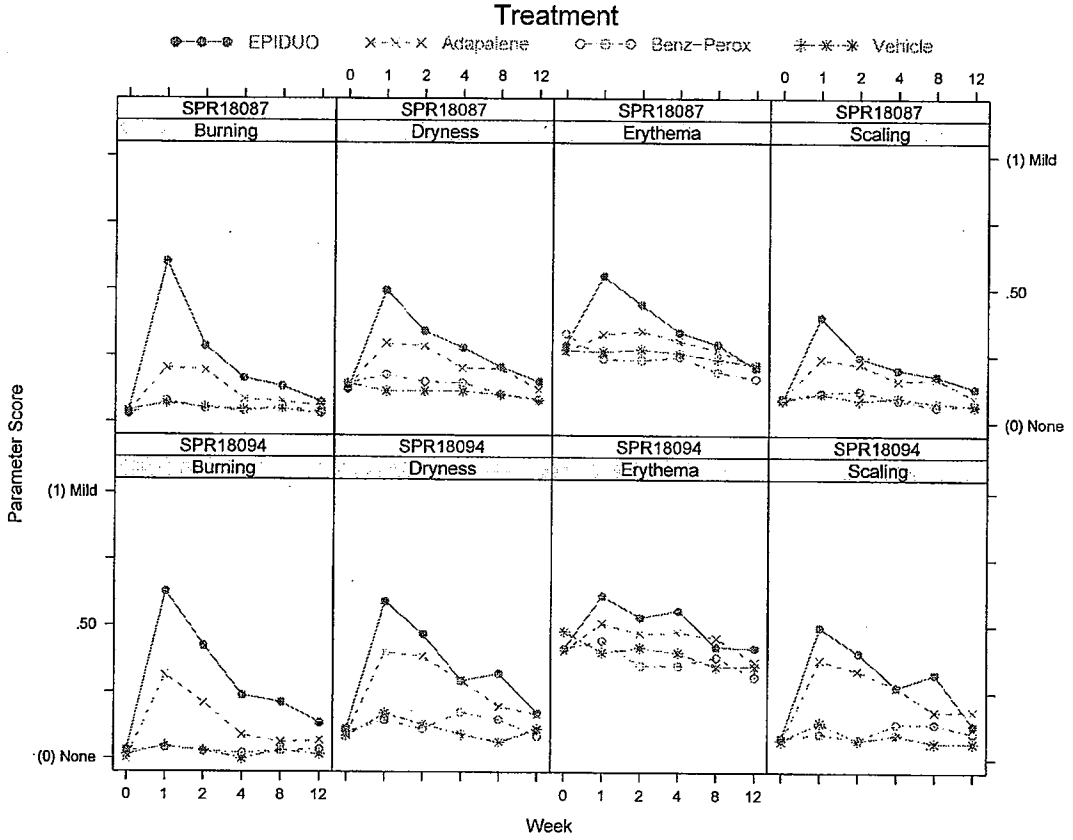
	EPIDUO ($N = 564$)	Adapalene ($N = 568$)	Benzoyl Peroxide ($N = 564$)	Vehicle ($N = 489$)
GENERAL DISORDERS AND ADMINISTRATION SITE CONDITIONS				
APPLICATION SITE BURNING	15 (2.7)	4 (0.7)	2 (0.4)	2 (0.4)
APPLICATION SITE IRRITATION	8 (1.4)	6 (1.1)	2 (0.4)	1 (0.2)
INFECTIONS AND INFESTATIONS				
NASOPHARYNGITIS	20 (3.5)	33 (5.8)	28 (5.0)	22 (4.5)
UPPER RESPIRATORY TRACT INFECTION	11 (2.0)	14 (2.5)	17 (3.0)	19 (3.9)
SINUSITIS	7 (1.2)	7 (1.2)	4 (0.7)	4 (0.8)
INJURY, POISONING AND PROCEDURAL COMPLICATIONS				
SUNBURN	7 (1.2)	9 (1.6)	3 (0.5)	5 (1.0)
NERVOUS SYSTEM DISORDERS				
HEADACHE	9 (1.6)	16 (2.8)	7 (1.2)	7 (1.4)
RESPIRATORY, THORACIC AND MEDIASTINAL DISORDERS				
PHARYNGOLARYNGEAL PAIN	7 (1.2)	6 (1.1)	7 (1.2)	5 (1.0)
COUGH	4 (0.7)	3 (0.5)	3 (0.5)	7 (1.4)
SKIN AND SUBCUTANEOUS TISSUE DISORDERS				
DRY SKIN	42 (7.4)	36 (6.3)	12 (2.1)	14 (2.9)
DERMATITIS CONTACT	18 (3.2)	20 (3.5)	4 (0.7)	3 (0.6)
PRURITUS	7 (1.2)	5 (0.9)	13 (2.3)	4 (0.8)

For subjects that experienced an AE multiple times, only a single instance is used in the tabulations below.
Source: Reviewer's Analysis of short-term safety from Study 18094 and Study 18087.

3.2.1.2 Local Skin Reactions Recall that four local skin reactions: buringing/stinging, dryness, erythema, and scaling were actively assessed at each visit with scores of 0 = 'None', 1 = 'Mild', 2 = 'Moderate', 3 = 'Severe'. The mean score over all subjects was calculated at each visit for each treatment group. Figure 13 depicts the mean profile over time for each study and local skin reaction grouped by treatment arm. For all local skin reactions the mean profile

of EPIDUO is above that of each monad and vehicle. During the twelve weeks of treatment the peak of the mean for each local reaction is at week 1 with a gradual reduction thereafter. Findings are consistent across the two studies.

Figure 13: Local Skin Reactions (Study 18087 and 18094)



3.2.2 Long-term Safety Evaluation

The safety data collected from Study 18089 resulted in 452 subjects being treated who were exposed to up to 12 months of treatment with EPIDUO. 361 (79.9%) subjects participated in the study for at least 6 months and 194 (42.9%) of subjects participated for at least 12 months.[8]

3.2.2.1 MedDRA Tabulation Table 29 contains the MedDRA preferred terms which were reported in at least 3.0% of subjects grouped according to the system organ classification (SOC).

[8]Study Report Table 13.

Comparing the long-term findings to the short-term findings, it can be seen that there is an increased percentage of subjects reporting an AE for application site burning, application site irritation, dry skin, erythema, and skin desquamation. These adverse events are related to the local skin reactions which are assessed further in the following section.

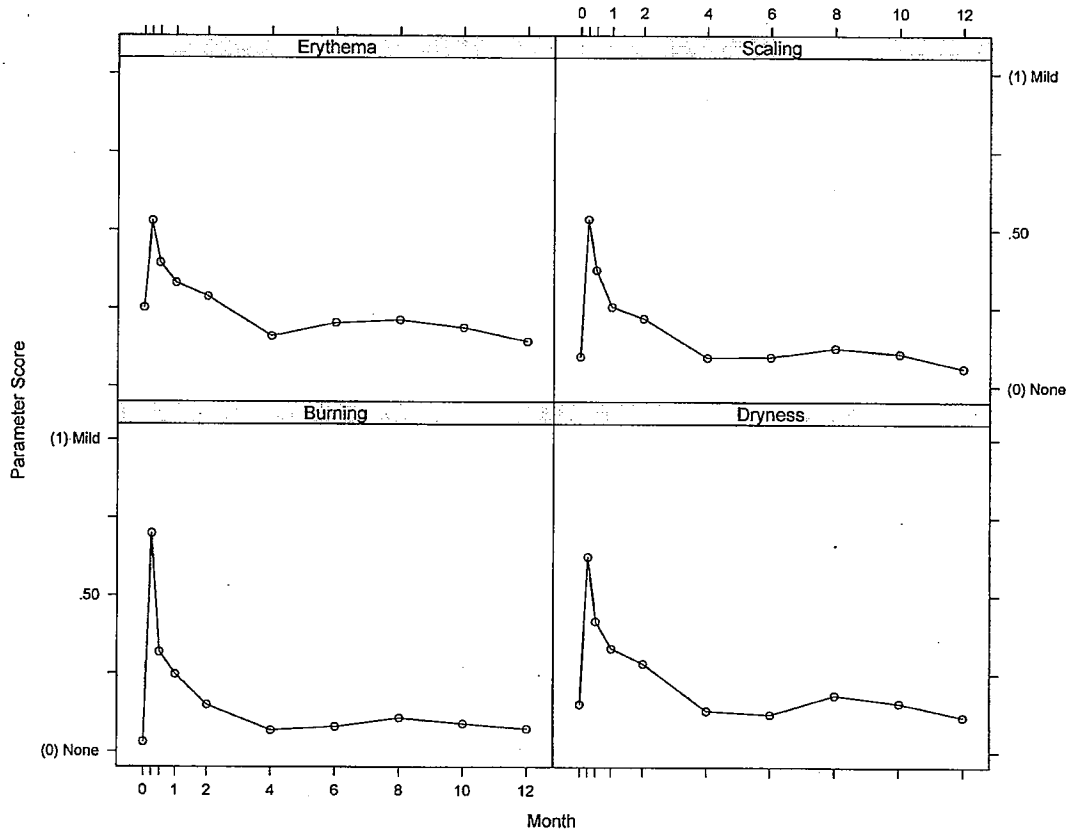
Table 29: Adverse Events by System Organ Class and Preferred Term

	EPIDUO ($N = 452$)
GENERAL DISORDERS AND ADMINISTRATION SITE CONDITIONS	
APPLICATION SITE BURNING	64 (14.2)
APPLICATION SITE IRRITATION	18 (4.0)
INFECTIONS AND INFESTATIONS	
NASOPHARYNGITIS	30 (6.6)
UPPER RESPIRATORY TRACT INFECTION	26 (5.8)
INFLUENZA	18 (4.0)
SINUSITIS	15 (3.3)
PHARYNGITIS STREPTOCOCCAL	13 (2.9)
INJURY, POISONING AND PROCEDURAL COMPLICATIONS	
SUNBURN	21 (4.6)
NERVOUS SYSTEM DISORDERS	
HEADACHE	21 (4.6)
RESPIRATORY, THORACIC AND MEDIASTINAL DISORDERS	
PHARYNGOLARYNGEAL PAIN	14 (3.1)
SKIN AND SUBCUTANEOUS TISSUE DISORDERS	
DRY SKIN	80 (17.7)
ERYTHEMA	25 (5.5)
SKIN DESQUAMATION	23 (5.1)
DERMATITIS CONTACT	17 (3.8)

For subjects that experienced an AE multiple times, only a single instance is used in the tabulations below. Source: Reviewer's Analysis of long-term safety from Study 18089.

3.2.2.2 Local Skin Reactions For each of the four local skin reactions the mean score over all subjects was calculated at each visit. Figure 14 depicts the mean profile over time for each of the four local skin reactions. As was seen in the short term studies the mean profile of EPIDUO is above that of each monad and vehicle. During the twelve weeks of treatment the peak of the mean for each local reaction is at week 1 with a gradual reduction thereafter.

Figure 14: Local Skin Reactions (Study 18089)



4 Findings in Special/Subgroup Populations

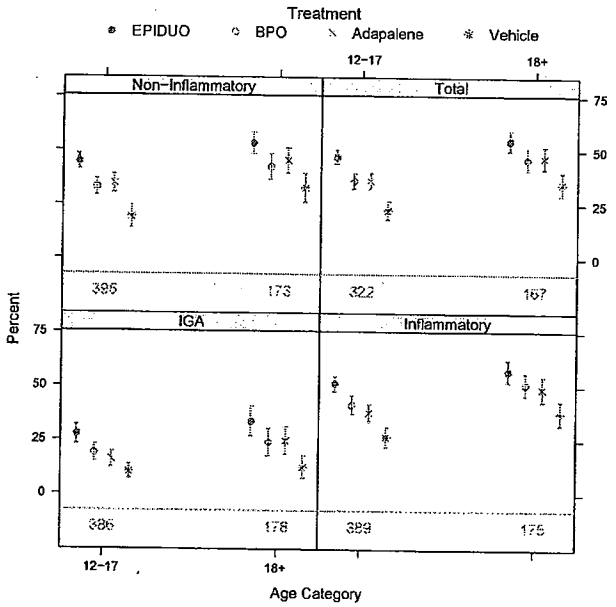
Section 4.1 provides a graphical assessment of efficacy by age, gender, and race. The data from Study 18094 and Study 18087 were combined so as to assess the general trend and patterns in the subgroups. Rather than using the absolute change from baseline in the graphical depictions, the percent change is used as this would be on a similar scale as the percent of IGA successes (IGA score of 'clear' or 'almost clear'). The analysis population is the ITT population with missing week 12 observations imputed using LOCF as in the primary analysis. Note that the protocol did not pre-specify any subgroup analysis which controlled the overall Type I error rate. For a tabular presentation of the data refer to the Appendix, Section A.2.1.

4.1 Gender, Race, and Age

4.1.1 Age

The age of subjects was dichotomized into two categories: 12 to 17 years old and 18 years and older. Figure 15 depicts efficacy results according to age category along with unadjusted 95% confidence intervals. The figure shows that EPIDUO had higher mean response rates than each of the monads and vehicle for both age groups for each of the endpoints. In general, subjects who were 18 and older tended to have slightly higher response rates than subjects 12 to 17 years old.

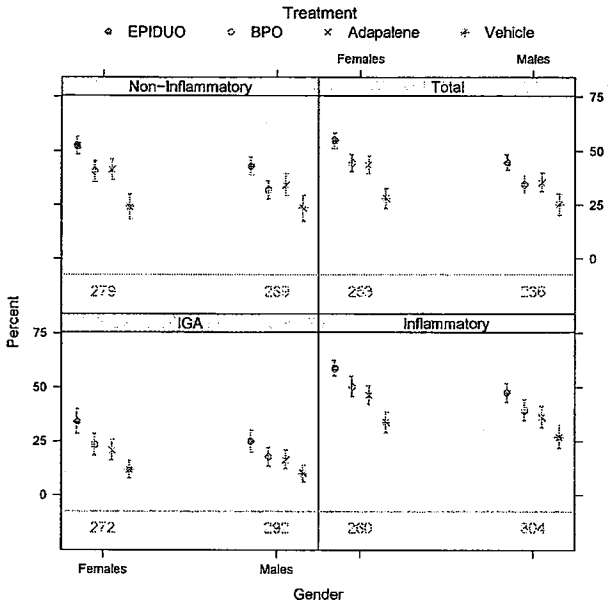
Figure 15: Efficacy Results According to Age



4.1.2 Gender

Figure 16 depicts efficacy results according to gender along with unadjusted 95% confidence intervals. In general EPIDUO had higher means than each of the monads and vehicle for both genders. For all endpoints, females tended to have higher means than males.

Figure 16: Efficacy Results According to Gender



4.1.3 Race

The number of subjects classified as Asian or Other is small from which to draw any definitive conclusions about these subgroups. Figure 17 depicts the means for each endpoint along with unadjusted 95% confidence intervals by race. Overall the efficacy results were quite consistent across subgroups.

4.2 Other Special/Subgroup Populations

4.2.1 Efficacy by Country

Study 18087 was conducted in 51 U.S centers, 6 Canadien centers, and 3 centers in Puerto Rico. Using the primary analysis population, Figure 18 depicts the mean percent of IGA successes for each country along with unadjusted 95% confidence intervals. Sample sizes for each treatment arm within a country are listed in the denominator of the figure. While sample sizes are smaller for both subjects enrolled in Puerto Rico and Canada, the mean response rates are higher in these two countries than in the U.S.

To provide further description of the efficacy results Table 30 contains IGA response rates for each treatment arm within country and a 95% confidence in the difference between EPIDUO and its monads and vehicle. As was also apparent in Figure 18, the treatment effect was larger when comparing EPIDUO to benzoyl peroxide than when comparing EPIDUO to adapalene in

Figure 17: Efficacy Results According to Race

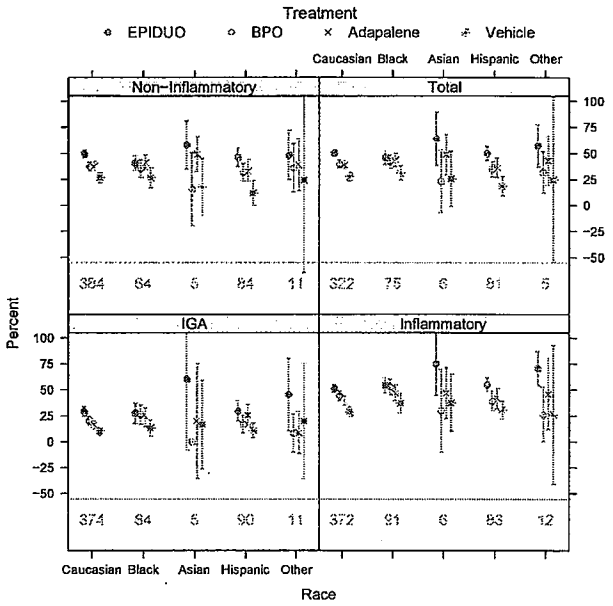
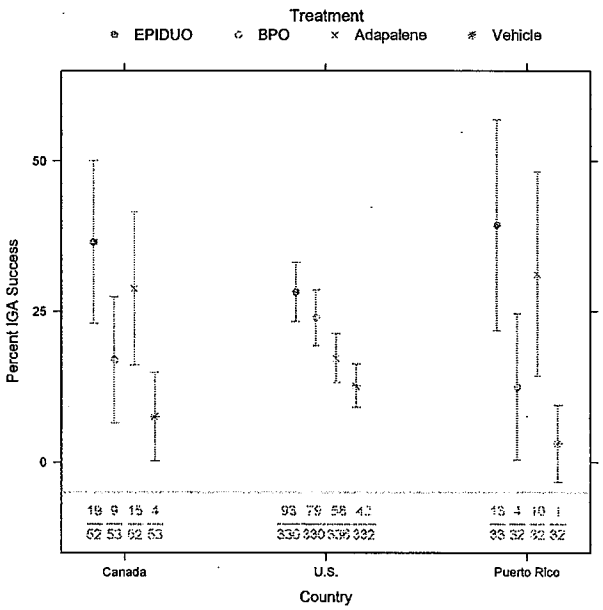


Figure 18: Efficacy Results According to Country (Study 18087)



both Canada and Puerto Rico. However, the opposite occurred in the U.S. An examination of demographic and prognostic factors between the countries did not reveal any differences from the U.S. population and the Canadien and Puerto Rican populations which would the explain the smaller response rate in U.S. subjects.

Table 30: Investigator Global Results (ITT-LOCF) by Country

	EPIDUO	Benzoyl Peroxide	Adapalene	Vehicle
Canada				
Percent x/N	36.5 19/52	17.0 9/53	28.8 15/52	7.5 4/53
δ (95% CI)	-	19.6 (1.1, 38)	7.7 (-12.2, 27.6)	29.0 (12.2, 45.8)
Puerto Rico				
Percent x/N	39.4 13/33	12.5 4/32	31.2 10/32	3.1 1/32
δ (95% CI)	-	26.9 (3.6, 50.2)	8.1 (-18.1, 34.4)	36.3 (15.5, 57.1)
U.S.A.				
Percent x/N	28.2 93/330	23.9 79/330	17.3 58/336	12.7 42/332
δ (95% CI)	-	4.2 (-2.8, 11.2)	10.9 (4.3, 17.5)	15.5 (9.2, 21.9)

Source: Reviewer's analysis.

5 Summary and Conclusions

5.1 Statistical Issues and Collective Evidence

Study 18094 was a Phase 2 trial. However in the submission of the protocol in SN00 the sponsor proposed that such a trial might be considered as one of two confirmatory Phase 3 trials. In response to SN00 it was noted that no formal statistical testing is required of Phase 2 trials, however general comments were provided to the sponsor about requirements of the statistical analysis for Phase 3 trials. It should be noted that none of these comments made after the Pre-IND meeting about the statistical analysis were incorporated into a revised protocol for Study 18094. In the review of the protocol for Study 18087, the sponsor and Division did reach agreements on the statistical analysis. For the assessment of efficacy, the agreements reached on the statistical analysis for Study 18087 were also applied to Study 18094. Refer to Section 2.2 for a description of the regulatory history and Section 3.1.2 for details of the statistical analysis agreed upon between the sponsor and the Division.

As Study 18094 enrolled subjects with baseline IGA scores of 'Mild' the assessment of the IGA dichotomized to success was performed using multiple definitions of success in Study 18094. For each of these definitions EPIDUO was consistently statistically superior to each of its monads

and vehicle (efficacy results are provided in Table 5 on page 20. In the analysis of the lesion counts, the treatment effects were larger in subjects with higher number of baseline lesions. At baseline it was detected that subjects randomized to EPIDUO had the lowest number of baseline lesions. Thus, as treatment effects were smaller for lower baseline lesion counts such an imbalance in the randomization did not provide a favorable condition to demonstrate efficacy of EPIDUO. Using a main effects model with terms for treatment, baseline lesion count, and site all contrasts comparing EPIDUO to its monads and vehicle were statistically significant at the $\alpha = 0.05$ level. Efficacy results for the lesion counts can be found in Tables 7 (inflammatory lesions), 8 (non-inflammatory lesions), and 9 (total lesions) on pages 23, 26, and 26, respectively.

Study 18087 was powered at over 90% using estimates of treatment effects from Study 18094. Based upon the pre-specified criteria, EPIDUO was superior to each of its monads and vehicle for the co-primary dichotomized IGA endpoint as well as for the change in non-inflammatory lesion endpoint. Efficacy results for these two endpoints are provided in Tables 17 and 19 on pages 34 and 38. For the co-primary endpoint, change in inflammatory lesion count, the comparison of EPIDUO to benzoyl peroxide did not reach statistical significance ($p = 0.068$) for the protocol defined primary analysis method. This analysis was based on ranking the data prior to fitting the ANCOVA model as ranks were used due to the significance of Shapiro-Wilk's normality test. An examination of the ANCOVA model diagnostics did not reveal an added benefit to using the ranks over the untransformed data in which case the analysis based on the untransformed data reached statistical significance ($p = 0.0387$). In addition, a sensitivity analysis using a transformation of the response, end of treatment inflammatory lesion count, with model terms for treatment, site, and baseline count showed more desirable properties in terms of assessing the model diagnostics. This sensitivity analysis also showed a statistically significant difference in inflammatory lesions between EPIDUO and benzoyl peroxide ($p = 0.0199$). Thus, the collective evidence suggests that EPIDUO is statistically superior to each of its monads and vehicle for *all* co-primary endpoints.

5.2 Conclusions and Recommendations

Study 18094 and Study 18087 were used to assess the efficacy of EPIDUO as compared to each monad and its vehicle. The primary efficacy endpoints were:

- Change from baseline in inflammatory and non-inflammatory lesion counts.
- Percent of patients with success ('clear' or 'almost clear') for the IGA.

In both studies, EPIDUO was statistically superior to each monad and vehicle for the percent of IGA successes and the change in non-inflammatory lesion counts. However, in Study 18087 for the co-primary endpoint, change in inflammatory lesion count, the comparison of EPIDUO to

benzoyl peroxide did not reach statistical significance ($p = 0.068$) for the protocol defined primary analysis method. However, several sensitivity analyses were conducted in which this comparison did reach the nominal $\alpha = 0.05$ level. In the safety assessment of local skin reactions, on average EPIDUO was more irritating than each monad and vehicle, especially within the first week of therapy. The mean intensity of the local skin reaction score for EPIDUO was below a mild rating, and the irritation tended to resolve and reach near baseline levels by week 12.

The following points relate to the original label proposed by the sponsor.

b(4)

b(4)

Therefore it is recommended that the above sentence be deleted from the label.

References

[1] Cleveland, W.S. *The Elements of Graphing Data.* Hobart Press, Summit, New Jersey, 1985.

[2] Statistical Analysis and Graphics produced with R software. R Development Core Team (2007). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL `http://www.R-project.org`.

APPENDIX

A.1 Supplementary Information for Study 18094, Study 18087, and Study 18089

A.1.1 Baseline Demographic Tables

Table 31: Demographic Factors by Treatment (Study 18094)

	EPIDUO $N = 149$	Adapalene $N = 148$	Benzoyl Peroxide $N = 149$	Vehicle $N = 71$
Age	14.0 15.0 17.0	14.0 16.0 17.0	14.0 16.0 17.0	14.0 15.0 18.5
Gender : Female	42% (62)	42% (62)	36% (53)	44% (31)
Race : Caucasian	68% (101)	70% (103)	77% (114)	73% (52)
Black	12% (18)	14% (20)	7% (10)	13% (9)
Asian	1% (1)	1% (1)	1% (2)	1% (1)
Hispanic	15% (23)	12% (18)	12% (18)	13% (9)
Other	4% (6)	4% (6)	3% (5)	0% (0)

a b c represent the lower quartile *a*, the median *b*, and the upper quartile *c* for continuous variables. Numbers after percents are frequencies.

Source: Study Report Table 9 and Reviewer Analysis.

Table 32: Demographic Factors by Treatment (Study 18087)

	EPIDUO $N = 415$	Adapalene $N = 420$	Benzoyl Peroxide $N = 415$	Vehicle $N = 418$
Age	14 16 20	14 16 19	14 16 20	14 16 19
Gender : Female	51% (210)	52% (217)	50% (207)	53% (222)
Race : Caucasian	66% (273)	67% (281)	62% (258)	65% (270)
Black	16% (66)	15% (64)	20% (81)	16% (66)
Asian	1% (4)	1% (4)	1% (4)	1% (5)
Hispanic	16% (67)	16% (66)	16% (65)	17% (72)
Other	1% (5)	1% (5)	2% (7)	1% (5)

a b c represent the lower quartile *a*, the median *b*, and the upper quartile *c* for continuous variables. Numbers after percents are frequencies.

Source: Study Report Table 6 and Reviewer Analysis.

Table 33: Baseline Demographic Factors (Study 18089)

	EPIDUO $N = 452$
Age	14.0 16.0 20.0
Gender : Female	51% (230)
Race : Caucasian	76% (345)
Black	12% (53)
Asian	2% (10)
Hispanic	7% (31)
Other	3% (13)

a b c represent the lower quartile *a*, the median *b*, and the upper quartile *c* for continuous variables. Numbers after percents are frequencies.

Source: Study Report Table 8 and Reviewer Analysis.

A.1.2 Baseline Prognostic Factors

Figure 19 depicts the empirical cumulative distribution (ECDF) plot for both lesion counts. If treatment groups had similar distributions of lesion counts the lines in the ECDF plot would be intersecting. However, this plot shows that the line for EPIDUO is consistently above the other treatment groups for lesion types implying subjects randomized to EPIDUO had fewer baseline lesions than the other treatment groups.

Figure 20 depicts the empirical cumulative distribution (ECDF) plot for both lesion counts. Similar to results seen in Study 18094, the plot shows that the line for EPIDUO is consistently above the other treatment groups for lesion types implying subjects randomized to EPIDUO had fewer baseline lesions than the other treatment groups. The baseline difference between treatment groups is smaller in Study 18087 than that observed in Study 18094.

Figure 19: Empirical Cumulative Distribution Plot (Study 18094)

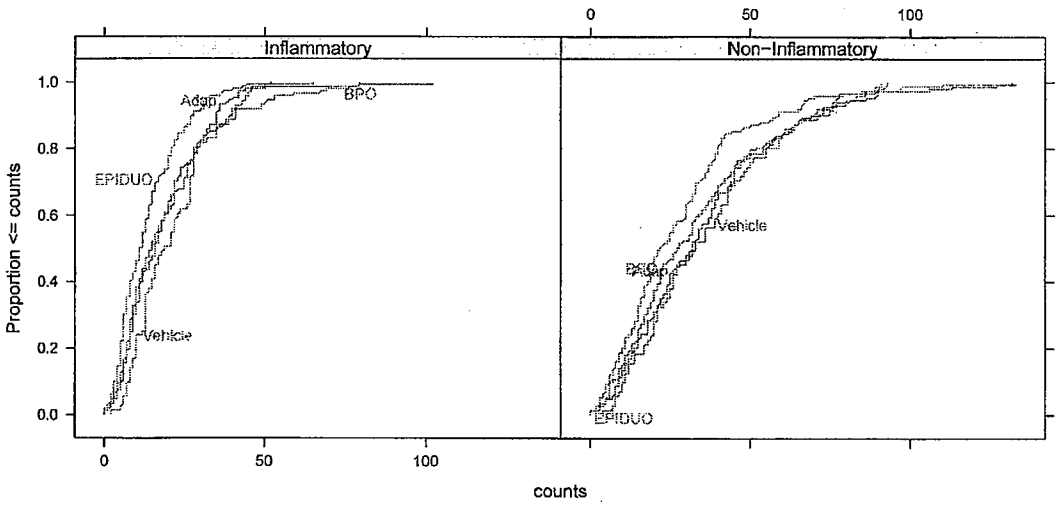
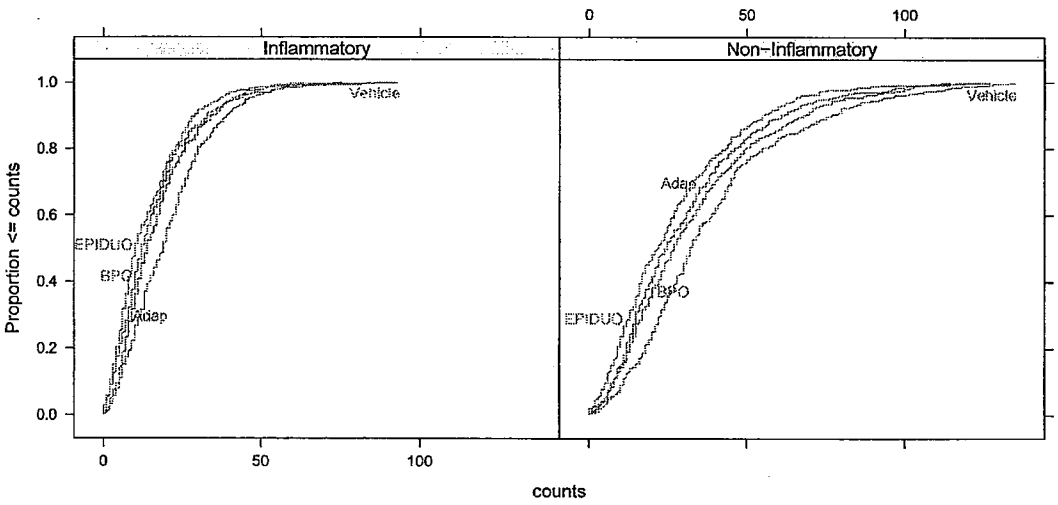


Figure 20: Empirical Cumulative Distribution Plot (Study 18087)



A.2 Diagnostic Plots on Ranked Data (ITT-LOCF)

The following figures show the diagnostic plots of the main effects model using a *rank* transform on the change in lesion count and the baseline lesion count. Both of the normal Q-Q plots show that the normality assumption of the residuals is not met. Note that there is also a slight violation of the constant variance assumption for the models in both lesion counts.

Figure 21: Model Diagnostics for Ranked Inflammatory Lesion Count

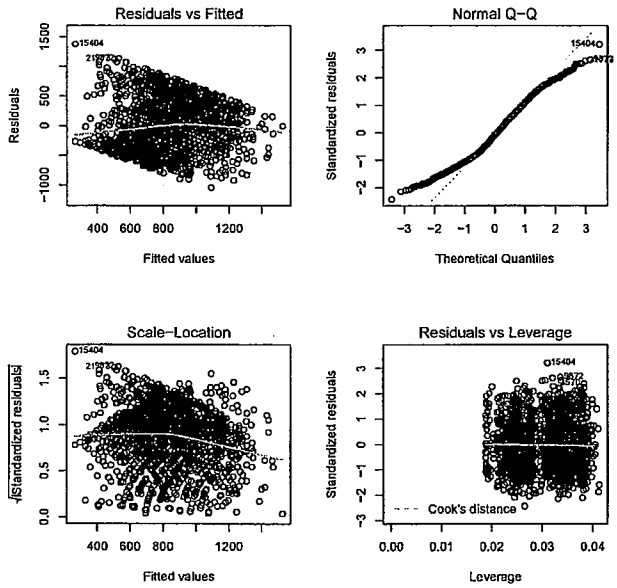
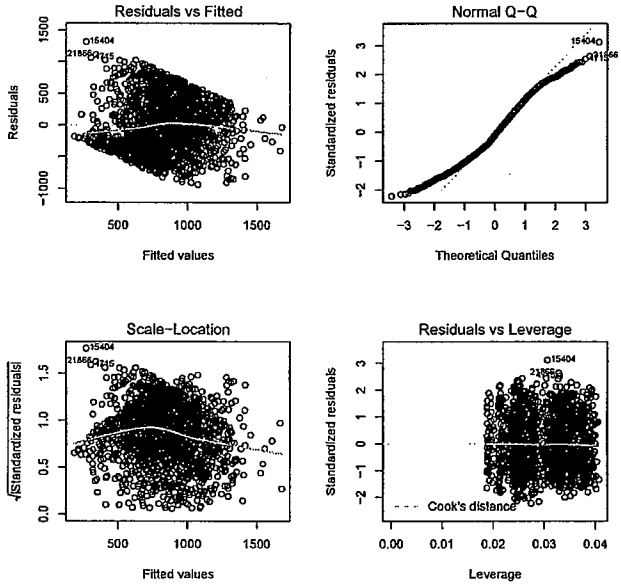


Figure 22: Model Diagnostics for Ranked Non-Inflammatory Lesion Count



A.2.1 Efficacy Tables by Subgroups

The following tables present point estimates of efficacy by age, gender, and race. In the reporting of lesion counts, the means and standard deviation are based upon absolute change from baseline whereas results presented in Section 4.1 were based upon the percent reduction from baseline.

A.2.1.1 Investigator Global Assessment Tables 34, 35, and 36 depict efficacy results for the endpoint: Investigator Global Assessment. Tabular information is separated out for each study with data shown being the percent successes (IGA score of 'clear' or 'almost clear') as well as fraction of successes for each subgroup. The analysis population was the ITT population with missing data imputed using LOCF.

Table 34: Investigator Global Results (ITT-LOCF) by Age

	EPIDUO	Benzoyl Peroxide	Adapalene	Vehicle
Study 18094				
12 to 17	28.1 $\frac{34}{121}$	13.8 $\frac{16}{116}$	15.5 $\frac{18}{116}$	4.2 $\frac{2}{48}$
18 and older	25 $\frac{7}{28}$	21.2 $\frac{7}{33}$	15.6 $\frac{5}{32}$	21.7 $\frac{5}{23}$
Study 18087				
12 to 17	27.2 $\frac{72}{265}$	20.9 $\frac{57}{273}$	16.1 $\frac{45}{279}$	11.4 $\frac{31}{273}$
18 and older	35.3 $\frac{53}{150}$	24.6 $\frac{35}{142}$	27 $\frac{38}{141}$	11.1 $\frac{16}{144}$

Source: Reviewer's analysis.

Table 35: Investigator Global Results (ITT-LOCF) by Gender

	EPIDUO	Benzoyl Peroxide	Adapalene	Vehicle
Study 18094				
Female	33.9 $\frac{21}{62}$	22.6 $\frac{12}{53}$	16.1 $\frac{10}{62}$	9.7 $\frac{3}{31}$
Male	23.0 $\frac{20}{87}$	11.5 $\frac{11}{96}$	15.1 $\frac{13}{86}$	10.0 $\frac{4}{40}$
Study 18087				
Female	34.3 $\frac{72}{210}$	23.7 $\frac{49}{207}$	22.1 $\frac{48}{217}$	12.2 $\frac{27}{221}$
Male	25.9 $\frac{53}{205}$	20.7 $\frac{43}{208}$	17.2 $\frac{35}{203}$	10.2 $\frac{20}{196}$

Source: Study Report Table 21; results reproduced by reviewer.

Table 36: Investigator Global Results (ITT-LOCF) by Race

	EPIDUO	Benzoyl Peroxide	Adapalene	Vehicle
Study 18094				
Caucasian	26.7 $\frac{27}{101}$	15.8 $\frac{18}{114}$	11.7 $\frac{12}{103}$	7.7 $\frac{4}{52}$
Black	33.3 $\frac{6}{18}$	20.0 $\frac{2}{10}$	25.0 $\frac{5}{20}$	11.1 $\frac{1}{9}$
Asian	100.0 $\frac{1}{1}$	0.0 $\frac{0}{2}$	0.0 $\frac{0}{1}$	0.0 $\frac{0}{1}$
Hispanic	26.1 $\frac{6}{23}$	16.7 $\frac{3}{18}$	33.3 $\frac{6}{18}$	22.2 $\frac{2}{9}$
Other	16.7 $\frac{1}{6}$	0.0 $\frac{0}{5}$	0.0 $\frac{0}{6}$	- $\frac{0}{0}$
Study 18087				
Caucasian	29.7 $\frac{81}{273}$	22.5 $\frac{58}{258}$	17.8 $\frac{50}{281}$	10.8 $\frac{29}{269}$
Black	25.8 $\frac{17}{66}$	27.2 $\frac{22}{81}$	23.4 $\frac{15}{64}$	13.6 $\frac{9}{66}$
Asian	50.0 $\frac{2}{4}$	0.0 $\frac{0}{4}$	25.0 $\frac{1}{4}$	20.0 $\frac{1}{5}$
Hispanic	31.3 $\frac{21}{67}$	16.9 $\frac{11}{65}$	24.2 $\frac{16}{66}$	9.7 $\frac{7}{72}$
Other	80.0 $\frac{4}{5}$	14.3 $\frac{1}{7}$	20.0 $\frac{1}{5}$	20.0 $\frac{1}{5}$

Source: Study Report Table 21; results reproduced by reviewer.

A.2.1.2 Change in Inflammatory Lesion Counts Tables 37, 38, and 39 depict efficacy results for the endpoint: absolute change from baseline for inflammatory lesions. Tabular information is separated out for each study with estimates for the mean and standard deviation (shown in parentheses). The analysis population was the ITT population with missing data imputed using LOCF. Cases where the statistic could not be calculated as no or only one subject was represented are denoted as 'NA'.

Table 37: Change in Inflammatory Lesions (ITT-LOCF) by Age Category

	EPIDUO	Benzoyl Peroxide	Adapalene	Vehicle
Study 18094				
12 to 17	-15.4 (12.4)	-9.9 (16.2)	-10.8 (10.8)	-7.8 (10.7)
18 and older	-18.4 (11.7)	-12.8 (11.1)	-13.3 (7.6)	-13.2 (11.0)
Study 18087				
12 to 17	-15.4 (11.5)	-13.5 (12.5)	-11.6 (13.6)	-8.0 (13.2)
18 and older	-15.4 (10.1)	-14.3 (10.3)	-13.7 (11.3)	-10.1 (10.7)

Numbers are mean and (standard deviation). Source: Reviewer's analysis.

Table 38: Change in Inflammatory Lesions (ITT-LOCF) by Gender

	EPIDUO	Benzoyl Peroxide	Adapalene	Vehicle
Study 18094				
Female	-18.2 (10.6)	-12.8 (13)	-11.5 (8.9)	-9.4 (9.3)
Male	-14.4 (13.1)	-9.3 (16.2)	-11.2 (11.1)	-9.7 (12.3)
Study 18087				
Female	-16.1 (9.6)	-14.1 (11.1)	-13.4 (11.1)	-9.5 (11.6)
Male	-14.6 (12.2)	-13.3 (12.4)	-11.1 (14.5)	-7.8 (13.3)

Numbers are mean and (standard deviation). Source: Reviewer's analysis.

Table 39: Change in Inflammatory Lesions (ITT-LOCF) by Race

	EPIDUO	Benzoyl Peroxide	Adapalene	Vehicle
Study 18094				
Caucasian	-15.2 (13.6)	-10.8 (15.8)	-10.7 (10.1)	-9.6 (10.9)
Black	-15.8 (10.7)	-12.4 (13.4)	-12.8 (11.0)	-6.0 (8.6)
Asian	-21.0 (NA)	-16.0 (1.4)	-13.0 (NA)	-3.0 (NA)
Hispanic	-18.7 (7.8)	-9.7 (12.9)	-14.9 (8.5)	-13.8 (13.7)
Other	-18.5 (7.3)	-1.0 (14.8)	-8.0 (14.2)	NA (NA)
Study 18087				
Caucasian	-19.5 (7.4)	-9.5 (16.8)	-13.8 (6.6)	-12.2 (8.6)
Black	-14.3 (9.8)	-14.9 (10.5)	-13.6 (11.6)	-9.7 (9.5)
Asian	-15.9 (11.7)	-14.1 (12.0)	-12.3 (13.1)	-8.5 (13.3)
Hispanic	-13.8 (9.5)	-11.5 (11.9)	-10.2 (13.7)	-8.3 (12.1)
Other	-18.4 (5.7)	-10.6 (10.8)	-20.4 (9.0)	-6.6 (13.3)

Numbers are mean and (standard deviation). Source: Reviewer's analysis.

A.2.1.3 Change in Non-Inflammatory Lesion Counts Tables 40, 41, and 42 depict efficacy results for the endpoint: absolute change from baseline for non-inflammatory lesions. Tabular information is separated out for each study with estimates for the mean and standard deviation (shown in parentheses). The analysis population was the ITT population with missing data imputed using LOCF. Cases where the statistic could not be calculated as no or only one subject was represented are denoted as 'NA'.

Table 40: Change in Non-Inflammatory Lesions (ITT-LOCF) by Age Category

	EPIDUO	Benzoyl Peroxide	Adapalene	Vehicle
Study 18094				
12 to 17	-22.7 (19.9)	-14.2 (16.4)	-14.8 (23.4)	-11.3 (19.7)
18 and older	-26.2 (20.1)	-12.1 (18.7)	-16.8 (18.1)	-17.1 (12.2)
Study 18087				
12 to 17	-23.9 (22.7)	-18.7 (22)	-20.6 (23.9)	-10.3 (28)
18 and older	-25.7 (19.2)	-20.3 (18.7)	-21.8 (19.8)	-13.2 (19.7)

Numbers are mean and (standard deviation). Source: Reviewer's analysis.

Table 41: Change in Non-Inflammatory Lesions (ITT-LOCF) by Gender

	EPIDUO	Benzoyl Peroxide	Adapalene	Vehicle
Study 18094				
Female	-27.5 (20.8)	-15.6 (16.6)	-15.1 (19.6)	-9.8 (17.2)
Male	-20.4 (18.8)	-12.6 (17)	-15.3 (24.3)	-15.8 (18)
Study 18087				
Female	-26.8 (20.8)	-19.3 (19)	-22.4 (23.2)	-11.8 (22.5)
Male	-22.3 (22.1)	-19.1 (22.7)	-19.5 (21.9)	-10.9 (28.5)

Numbers are mean and (standard deviation). Source: Reviewer's analysis.

Table 42: Change in Non-Inflammatory Lesions (ITT-LOCF) by Race

	EPIDUO	Benzoyl Peroxide	Adapalene	Vehicle
Study 18094				
Caucasian	-23.3 (18.7)	-13.5 (17.6)	-14.9 (23.9)	-12.6 (17.5)
Black	-22.4 (24.4)	-14.7 (8.8)	-14.5 (19.1)	-13.7 (19.3)
Asian	-21 (NA)	-1.5 (14.8)	-25 (NA)	15 (NA)
Hispanic	-26.2 (22.7)	-14.9 (17.1)	-18.2 (20.3)	-19.3 (16.9)
Other	-17 (20.6)	-17.6 (15.1)	-12.7 (13.2)	NA (NA)
Study 18087				
Caucasian	-31.5 (13.5)	-12.8 (18.9)	-23.2 (13.7)	-10 (7.1)
Black	-21.9 (19.9)	-17.5 (21)	-23.7 (19)	-15 (21.2)
Asian	-24.4 (20.2)	-20.1 (20.2)	-21.3 (20.8)	-12.7 (22.5)
Hispanic	-26.7 (27.5)	-18.3 (24)	-16.6 (32)	-3.1 (36.6)
Other	-31.8 (29.5)	-16.9 (18.3)	-27.2 (20.8)	-10.2 (28.6)

Numbers are mean and (standard deviation). Source: Reviewer's analysis.

Signatures/Distribution List

Primary Statistical Reviewer: Mat Soukup, Ph.D.
Date: October 21, 2008

Statistical Team Leader: Mohamed Alosh, Ph.D.

cc:
Archival NDA
DDDP/Walker
DDDP/Lindstrom
DDDP/Liedtka
DDDP/Williams
OBIO/Patrician
DB3/Wilson
DB3/Alosh
DB3/Soukup

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

/s/

Matt Soukup
10/21/2008 09:51:41 AM
BIOMETRICS

Mohamed Alosh
10/21/2008 10:13:49 AM
BIOMETRICS

Statistical Review and Evaluation
Fileability Review

NDA Number:	22-320
Drug Name:	EPIDUO (adapalene 0.1%/benzoyl peroxide 2.5%) Gel
Applicant:	Galderma
Indication:	Acne vulgaris
Filing Date:	04/08/2008
Fileability Meeting Date:	03/19/2008
User Fee Date:	12/08/2008
Received for Stat Review:	02/14/2008
Statistical Reviewer:	Mat Soukup, Ph.D., DBIII
Medical Officer:	Jane Liedtka, M.D., DDDP
Project Manager:	Kalyani Bhatt, DDDP

1 Background and Summary

This is a 505(b)(2) application submitted in electronic form. The proposed drug product contains two active drug substances, adapalene 0.1% and benzoyl peroxide 2.5% in a gel vehicle. The sponsor is seeking the indication of acne vulgaris. The clinical studies were conducted under IND 67,801.

Study SRE.2674 was a dose-finding study to confirm the best tolerated dose of benzoyl peroxide when combined with adapalene 0.1% in the fixed-combination product. Two pivotal trials were conducted in subjects 12 years and older. These studies were SRE.18094 and SRE.18087 which consisted of the combination product, the two monads, and vehicle. Treatment was once daily for 12 weeks.

2 Organization and Data Representation

1. Is there a comprehensive table of contents with adequate indexing and pagination? *Yes*
2. Are the original protocols, protocol amendments, and proposed label provided? *Yes. Protocols are located in each study report, and the label is available in EDR.*
3. Based on either the electronic data sets or the study reports can the following information be reviewed?
   (a) Patient profile listings by center for all enrolled subjects. *Yes, this will be possible with the electronic data sets.*
   (b) Discontinued subject tables by center (includes reason and time of loss). *Yes, the information is available in the data set SUB180xx using variables TERMIN and EXIT.*

(c) Subgroup analysis summary tables (gender, race, age, etc.). *Yes, both study reports and electronic data set: SUB180xx.*

(d) Adverse event listings by center and time of occurrence. *Yes, this is available in the AEF180xx data set; note that AE events are reported using COSTART terminology.*

4. Information specific to the electronically submitted data.

(a) Has adequate documentation of the data sets been provided? *No. There is insufficient information available defining the visit schedule. The analysis data set uses a derived variable called XVISIT which corresponds to results as presented in the study reports. However, there is no date of visit included in the efficacy data set nor is there a description of how XVISIT was derived. In addition, XVISIT is not included in all data sets which does not allow merging of data sets based on both the unique subject ID and the visit.*

(b) Do the data appear to accurately represent the data described in the study reports? *Based on the analysis data set and using the derived variable XVISIT efficacy results in the study reports could be reproduced. However, it is unclear how the variable XVISIT was derived. In addition, demographics and baseline characteristics could be reproduced.*

(c) Can the data be easily merged across studies and indications? *No (see above).*

3 Statistical Methodology

1. Are all primary efficacy studies of appropriate design to meet basic approvability requirements within current Division policy or to the extent agreed upon previously with the sponsor by the Division. *Statistical analyses appear to coincide with agreements/comments given during the review of Protocol 18087.*

2. For each study, is there a comprehensive statistical summary of the efficacy which covers the intent-to-treat population and per protocol population? *Yes.*

3. Based on the summary analyses of each study:

(a) Are the analyses appropriate for the type of data collected, the study design, and the study objectives (based on protocol objectives and proposed labeling claims)? *Yes.*

(b) Are the intent-to-treat and per protocol patient analyses properly performed? *Yes, it appears to be adequate.*

(c) Has missing data been appropriately handled? *Yes, this is LOCF. The sponsor has also conducted sensitivity analyses which appear to be in agreement with the Division's previous comments.*

(d) Have multiplicity issues (regarding endpoints, timepoints, or dose groups) been adequately addressed? *N/A*

(e) If interim analyses were performed, were they planned in the protocol and appropriate significance level adjustments made? *N/A*

4. Were sufficient and appropriate references included for novel statistical approaches? *N/A*

5. Are all pivotal studies complete? *Yes.*

6. Has the safety data been comprehensively and adequately summarized? *Yes, this appears to be the case based upon the study reports.*

4 Fileability Conclusions

From a statistical perspective this submission, or indications therein, *is not* reviewable without further input from the sponsor at this time.

5 74-Day Letter Comments

Issue: Efficacy results as presented in the study report for Studies 18087 and 18094 appear to rely on the use of the derived analysis visit variable `XVISIT`. However, the sponsor has not included details of how this analysis visit variable was defined. Further, this analysis visit is not included in all data sets which does not allow merging of data sets by unique subject ID and visit. Consequently, this does not allow for a thorough review of the data submitted in the NDA.

Request for Information: The sponsor is requested to submit new electronic data sets which include a consistent definition of the analysis visit allowing for merging of data sets by unique subject ID *AND* visit. Any derived analysis visit should also include a detailed description of how such a variable was derived. In addition, it is requested that the date of the visit also be included in the electronic data. The sponsor should note that results as presented in the study reports should be reproducible using the submitted data sets.

Mat Soukup, Ph.D.
Mathematical Statistician, Biometrics 3

Concur: Mohamed Alosh, Ph.D.
Team Leader, Biometrics 3

Cc:
Orig. NDA 22,320/SN000
DDDP/Walker
DDDP/Lindstrom
DDDP/Liedtka
DDDP/Bhatt
OBIO/Patrician
DBIII/Wilson
DBIII/Alosh
DBIII/Soukup

April 4, 2008

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

/s/

Matt Soukup
4/4/2008 02:25:04 PM
BIOMETRICS

Mohamed Alosh
4/4/2008 02:55:28 PM
BIOMETRICS