CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER: 20-508/S005

# STATISTICAL REVIEW(S)

540/White

JUL 2 4 2000

# Statistical Review and Evaluation

**NDA:** 20-508

**Name of Drug:** Lac-Hydrin (ammonium lactate) Cream, 12%

**Applicant:** Westwood Squibb Pharmaceuticals Inc.
100 Forest Avenue
Buffalo, New York 14213-1091

**Indications:** Ichthyosis vulgaris and xerosis in children

**Documents Reviewed:** Volumes 1-5 of submission withd diskettes containing SAS data sets from the sponsor

**Medical Officer:** Denise Cook, M.D. (HFD-540)

## Introduction:

The goal of this Phase IV study was to obtain information on local tolerance and clinical safety of Lac-Hydrin Cream. 12%, in children, to support labeling for pediatric use in the treatment of ichthyosis vulgaris and xerosis. Since the incidence of skin related adverse events was felt to be higher in patients with ichthyosis vulgaris, that was the condition used for entry into the study.

## Design (Protocol DE109-035):

This was a randomized, double-blind, vehicle-controlled, parallel group study of Lac-Hydrin Cream, 12%, and its vehicle in the treatment of ichthyosis vulgaris. Treatment was applied twice daily to all affected areas. After a four week treatment period there was a two week no treatment wash-out period to examine the persistence of treatment effects. Some 103 subjects were enrolled in 12 centers for this study. Three efficacy measures: overall severity, severity of scaling, and severity of fissuring, all on an 9-point scale, were used:

| Score | Overall Severity | Scaling | Fissuring |
|---|---|---|---|
| 0 | No evidence | No evidence | No evidence |
| 1 | | Fine scaling, limited distribution | |
| 2 | Mild | Fine scaling , wide distribution | Fine, limited fissuring |
| 3 | | Faint,distinct polygonal scales | |
| 4 | Moderate | Distinct polygonal scales | Moderate fissuring |
| 5 | | Moderate number of polygonal scales | |
| 6 | Severe | Large number of polygonal scales | Distinct areas of fissuring |
| 7 | | Most area covered with thick plates | |
| 8 | Very Severe | Completely covered by thick plates | Severe fissuring |

For the efficacy analysis the primary endpoint was the overall disease severity, with the others treated as secondary. But note that for this particular study the safety information was considered paramount.

It was the opinion of the Medical Officer that for some subjects the amount of body surface area involved (treated) was considered to be inconsistent with the diagnosis of icthyosis vulgaris. These were six patients: patient 105 (Lac-Hydrin), 121(Lac-Hydrin), 122 (Lac-Hydrin), 123 (vehicle), 126 (Lac-Hydrin), and 127 (vehicle). It was felt that these patients should be deleted from the efficacy analysis. Since this deletion was defined post-randomization it defines a "Modified" Intent to Treat (MITT) Group. For safety all subjects were analyzed.

## Statistical Methods:

In the sponsor's analysis of this study "Primary and secondary efficacy measures were analyzed using the investigator-adjusted Wilcoxon rank-sum test to evaluate the null hypothesis of no treatment difference between treatment groups. Treatment differences in the proportion of subjects experiencing at least one skin-related adverse event were evaluated by investigator-adjusted Cochran Mantel-Haenszel test for general association to test the hypothesis of no difference between treatment groups." (Page 4 of Final Report)

Note that the Wilcoxon rank-sum (i.e. Mann-Whitney) test is actually a test that the medians are equal in populations with similar dispersion. The sponsor seems to use this to explain differences in means. For symmetric populations the mean and median will be equal, but otherwise they can differ. In fact, at most time points the observed data were reasonably symmetric, particularly in the primary endpoint, with roughly comparable dispersions across groups. So the analysis specified by the sponsor seems appropriate. However, it is this reviewer's opinion that for such studies analysis of variance is generally quite robust, and a priori. would have provided a somewhat preferable analysis (possibly more powerful, but at least more directly related to the means). But since the rank test above was specified by the study protocol. it was also used in this reviewer's analyses.

The sponsor provides one primary endpoint and two secondary endpoints for efficacy. each measured at seven time points. The natural endpoint for such a study is the end of treatment (EOT, i.e.. Visit 6, Week 4), and was specified in the protocol as the principal evaluation time point. Tests at other time points are provided only as rough guides, and multiplicity issues attendant to multiple measures across time are ignored. With only two endpoints it is this reviewer's opinion that a reasonable argument can be made for also ignoring the multiplicity issues associated with the analysis of more than one endpoint.

In the first paragraph above, it was noted that the sponsor proposed to use a CMH test, stratified on investigator, to compare the incidence of various types of adverse events. Since there are relatively many centers compared to the number of adverse events, some centers will likely have a marginal total of zero adverse events. Centers with such zero marginal totals are completely dropped from the computation of the CMH test statistic. Under such circumstances the proposed CMH test may not be appropriate. If skin-related adverse events are relatively rare, as seems to occur here, necessarily several centers will have no adverse events and hence will be dropped from the analysis. Under such circumstances it would seem that a chi-square test comparing incidence of adverse events, or perhaps better, a Fisher Exact test, ignoring stratification on investigator would seem to be more appropriate.

## Efficacy Results:

As with the sponsor's original analysis, during weeks 0 to 4 (end of treatment), missing scores are imputed using "last observation carried forward" (LOCF) technology. Then the mean overall severity score at Visit 6 (end of treatment) for subjects using Lac-Hydrin 12% Cream was 1.9 compared with 3.1 for subjects using vehicle cream. The mean overall severity score with Lac-Hydrin 12% Cream was statistically significantly lower ($p \leq 0.0012$) than vehicle cream at Visit 6 (Week 4). The mean overall severity score with Lac-Hydrin 12% Cream was also statistically significantly lower ($p \leq 0.0194$) than vehicle cream at Visit 5 (Week 3). A week after treatment was stopped (Visit 7), overall severity was still statistically significantly lower ($p \leq 0.0307$) in the Lac-Hydrin 12% Cream treatment group. At Visit 8, two weeks after treatment was stopped, no significant difference in overall severity was observed between the two treatment groups. Mean and median overall severity scores by visit for the Modified Intent-to-Treat population (with data imputation using LOCF) are shown in Table 1, below. The mean is shown as the second entry in each cell, below the cell count, with the corresponding median in parentheses.

**Table 1. Summary of Overall Severity (Modified Intent-to-Treat)**

| Visit | Overall Severity | Lac-Hydrin 12% Cream (N=48) | Vehicle Cream (N=49) | p-value* |
|---|---|---|---|---|
| Baseline (Week 0) | n<br>mean ( median) | 48<br>5.6 (6) | 49<br>5.4 (5) | 0.3054 |
| Visit 3 (Week 1) | n<br>mean ( median) | 48<br>4.1 (4) | 49<br>4.1 (4) | 0.9648 |
| Visit 4 (Week 2) | n<br>mean ( median) | 48<br>3.1 (3) | 49<br>3.6 (4) | 0.2360 |
| Visit 5 (Week 3) | n<br>mean ( median) | 48<br>2.3 (2) | 49<br>3.2 (3) | 0.0194 |
| **Visit 6 –EOT (Week 4)** | **n<br>mean ( median)** | **48<br>1.9 (2)** | **49<br>3.1 (3)** | **0.0012** |
| Visit 7 (Week 5) | n<br>mean ( median) | 43<br>2.7 (3) | 42<br>3.6 (4) | 0.0307 |
| Visit 8 (Week 6) | n<br>mean ( median) | 46<br>4.0 (4) | 43<br>4.1 (4) | 0.9867 |

*p-value based on the Wilcoxon rank-sum test.

Tables 2 and 3 below, provide similar results for the secondary endpoints scaling and fissuring. From table 2 one can see that the mean scaling severity at Visit 6 (end of treatment) for subjects using Lac-Hydrin 12% Cream was 1.7 compared with 3.1 for subjects using vehicle cream. The mean scaling severity score with Lac-Hydrin 12% Cream was statistically significantly lower ($p \leq 0.0001$) than vehicle cream at this Visit 6 (Week 4). A week after treatment was stopped (Visit 7), scaling was still significantly lower ($p \leq 0.011$) in the Lac-Hydrin 12% Cream treatment group. At Visit 8, two weeks after treatment was stopped, there was no statistically significant difference between the two treatment groups in scaling.

**Table 2. Summary of Scaling (Modified Intent-to-Treat)**

| Visit | Scaling | Lac-Hydrin 12% Cream (N=48) | Vehicle Cream (N=49) | p-value* |
|---|---|---|---|---|
| Baseline (Week 0) | n<br>mean ( median) | 48<br>5.7 (6) | 49<br>5.5 (6) | 0.4247 |
| Visit 3 (Week 1) | n<br>mean ( median) | 48<br>4.0 (4) | 49<br>4.0 (4) | 0.8484 |
| Visit 4 (Week 2) | n<br>mean ( median) | 48<br>3.1 (3) | 49<br>3.4 (3) | 0.4580 |
| Visit 5 (Week 3) | n<br>mean ( median) | 48<br>2.2 (2) | 49<br>2.9 (3) | 0.0516 |
| **Visit 6 –EOT (Week 4)** | **n<br>mean ( median)** | **48<br>1.7 (1)** | **49<br>3.1 (3)** | **0.0001** |
| Visit 7 (Week 5) | n<br>mean (median) | 43<br>2.5 (2.5) | 42<br>3.6 (4) | 0.0110 |
| Visit 8 (Week 6) | n<br>mean (median) | 46<br>4.0 (4) | 43<br>4.1 (4) | 0.7898 |

*p-value based on the Wilcoxon rank-sum test.

Table 3, below, gives similar results for fissuring. At Visit 6 (end of treatment) for subjects using Lac-Hydrin 12% Cream the mean was 1.3 compared with 2.2 for subjects using vehicle cream. The mean fissuring severity score with Lac-Hydrin 12% Cream was statistically significantly lower ($p \leq 0.0024$) than vehicle cream at Visit 6 (Week 4). The mean fissuring score with Lac-Hydrin 12% Cream was also statistically significantly lower ($p \leq 0.0124$) than vehicle cream at Visit 5 (Week 3). A week after treatment was stopped (Visit 7), fissuring severity was still, though barely, statistically significantly lower ($p \leq 0.0453$) in the Lac-Hydrin 12% Cream treatment group. As with the other measures, by Visit 8, two weeks after treatment was stopped, no significant differences were observed.

**Table 3. Summary of Fissuring (Modified Intent-to-Treat)**

| Visit | Fissuring | Lac-Hydrin 12% Cream (N=48) | Vehicle Cream (N=49) | p-value* |
|---|---|---|---|---|
| Baseline (Week 0) | n<br>mean (median) | 48<br>4.4 (5) | 49<br>4.4 (5) | 0.9912 |
| Visit 3 (Week 1) | n<br>mean (median) | 48<br>2.7 (3) | 49<br>2.8 (3) | 0.7259 |
| Visit 4 (Week 2) | n<br>mean (median) | 48<br>2.1 (2) | 49<br>2.4 (2) | 0.3279 |
| Visit 5 (Week 3) | n<br>mean (median) | 48<br>1.3 (0) | 49<br>2.0 (2) | 0.0124 |
| **Visit 6 –EOT (Week 4)** | **n<br>mean (median)** | **48<br>1.3 (0)** | **49<br>2.2 (2)** | **0.0024** |
| Visit 7 (Week 5) | n<br>mean (median) | 43<br>1.7 (1) | 42<br>2.6 (2) | 0.0453 |
| Visit 8 (Week 6) | n<br>mean (median) | 46<br>3.2 (3) | 43<br>3.0 (3) | 0.4809 |

*p-value based on the Wilcoxon rank-sum test.

The following table provides a breakdown of the effect of various demographic groupings on the week 6 overall severity score. These tests are not powered to detect differences in these groups, but clearly there is no statistically significant evidence of a difference across race ($p \leq 0.9854$) or across the three age groups ($p \leq 0.3238$). Differences across gender seem somewhat more substantial in magnitude, but are still not statistically significant ($p \leq 0.1853$). Deciding whether or not this potential effect is just an artifact of the study or is worthy of further consideration is a decision requiring the expertise of the Medical Officer.

**Table 4. Summary of Overall Severity of Ichthyosis Vulgaris by Gender, Race, and Age Group at Visit 6 (Modified Intent-to-Treat)**

| Subgroups | Lac-Hydrin 12% Cream | | Vehicle Cream | | |
|---|---|---|---|---|---|
| | n | Mean | n | Mean | p-value |
| Gender* | | | | | 0.1853 |
| Male | 33 | 2.1 | 37 | 3.2 | |
| Female | 15 | 1.4 | 12 | 2.8 | |
| Race* | | | | | 0.9854 |
| White | 35 | 1.9 | 36 | 3.1 | |
| Non-White | 13 | 1.9 | 13 | 2.9 | |
| Age Group* | | | | | 0.3238 |
| 2 – 6 years | 11 | 1.5 | 16 | 2.7 | |
| 7 – 11 years | 21 | 1.9 | 18 | 3.1 | |
| 12 – 16 years | 16 | 2.1 | 15 | 3.5 | |

*Analyzed using two-way ANOVA with the rank of Visit 6 overall severity as the dependent variable.

Note the above analyses, as reflected in tables 1 through 4, all use the MITT population specified by the Medical Officer. But the conclusions remain consistent with those of the sponsor.

## Safety Results:

For this supplement, the safety results may be of more importance than the efficacy analysis above. However, it was the opinion of the Medical Officer that the appropriate study population was all subjects allocated medication, and no further statistical analysis was needed beyond that provided by the sponsor. Thus this reviewer's analysis was primarily limited to confirming the outcomes reported by the sponsor in the computer data sets provided to this reviewer.

Note that 52 adverse events (AEs) were reported by 36 subjects. Eighteen subjects in each treatment group experienced at least one adverse event. Most of these were of mild intensity and were claimed by the sponsor to be unrelated to treatment.

Fourteen percent (7/52) of the subjects in the Lac-Hydrin 12% Cream group and 10% (5/51) in the vehicle cream group had at least one nominally treatment related AE. All these AEs involved the "skin/appendages" body group. The CMH test, provided by the sponsor for testing differences across treatment group, was not statistically significant ($p \leq 0.3555$). The significance

level of the arguably more appropriate Fisher exact test, ignoring stratification on centers, was also not statistically significant ($p \leq 0.760$).

Twenty three percent (12/52) subjects reported skin-related AEs (whether treatment related or not) in the Lac-Hydrin 12% Cream group versus 14% (7/51) in the vehicle group. The significance level of the Fisher exact test, comparing the occurrence of skin-related AEs across treatment was also not statistically significant ($p \leq 0.310$). The most common skin-related adverse event was "burning skin", reported in 10% of the subjects in the Lac-Hydrin 12% Cream group versus 6% in the vehicle cream group ($p \leq 0.715$ from Fisher Exact test).

Again, the lack of statistical significance can be explained by the relatively small sample sizes. Note that there is a close to 2-1 ratio of subjects reporting skin-related adverse events or report "burning skin". However the study was not designed to test these effects, so deciding whether or not these ratios are of clinical importance requires the expertise of the Medical Officer.

APPEARS THIS WAY
ON ORIGINAL

## Conclusions:

1. The goal of this Phase IV study was to obtain information on local tolerance and clinical safety of Lac-Hydrin Cream, 12%, in children, to support labeling for pediatric use in the treatment of ichthyosis vulgaris and xerosis.

2. For the efficacy analysis three endpoints were evaluated: overall severity, scaling, and fissuring, each measured on a 9 point scale, 0-8, with smaller values reflecting less severity. The primary endpoint was the overall disease severity, with the others treated as secondary. But note that for this particular study the safety information was considered paramount.

3. It was the opinion of the Medical Officer that for some subjects the amount of body surface area involved (treated) was considered to be inconsistent with the diagnosis of icthyosis vulgaris. There were six such patients: four in the Lac-Hydrin Crean, 12% group, and two in the vehicle group. These patients were deleted from the efficacy analysis. Since this deletion was defined post-randomization, it defines a "Modified" Intent to Treat (MITT) Group for efficacy analysis. For safety all subjects were analyzed.

4. The mean overall severity score at Visit 6 (week 4, end of treatment) for subjects using Lac-Hydrin 12% Cream was 1.9 compared with 3.1 for subjects using vehicle cream ($p \leq 0.0012$). The mean overall severity score with Lac-Hydrin 12% Cream was also statistically significantly lower ($p \leq 0.0194$) than vehicle cream at Visit 5 (Week 3), and at Visit 7, a week after treatment was stopped ($p \leq 0.0307$). At Visit 8, two weeks after treatment was stopped, no statistically significant difference in overall severity was observed between the two treatment groups.

5. Fourteen percent (7/52) of the subjects in the Lac-Hydrin 12% Cream group and 10% (5/51) in the vehicle cream group had at least one nominally treatment related adverse event (AE). All these AEs involved the "skin/appendages" body group. The CMH test, provided by the sponsor for testing differences across treatment group, was not statistically significant ($p \leq 0.3555$). The significance level of the possibly more appropriate Fisher exact test, ignoring stratification on centers, was also not statistically significant ($p \leq 0.760$).

6. Twenty three percent (12/52) subjects reported skin-related AEs (whether treatment related or not) in the Lac-Hydrin 12% Cream group versus 14% (7/51) in the vehicle group. The significance level of the Fisher exact test, comparing the occurrence of skin-related AEs across treatment was not statistically significant ($p \leq 0.310$). The most common skin-related adverse event was "burning skin", reported in 10% of the subjects in the Lac-Hydrin 12% Cream group versus 6% in the vehicle cream group ($p \leq 0.715$ from Fisher Exact test).

7. The lack of statistical significance in tests comparing adverse events can be explained by the relatively small sample sizes. Since the study was not designed to test such effects, deciding whether or not the close to 2-1 ratio of subjects reporting skin-related adverse events or reporting "burning skin" is of any clinical relevance is a decision requiring the expertise of the Medical Officer.

APPEARS THIS WAY
ON ORIGINAL

/S/ 7/24/00

**Steve Thomson**
**Mathematical Statistician, Biometrics III**

/S/ 7/24/00

concur: **Mohamed Al Osh, Ph.D.**
**Acting Team Leader, Biometrics III**

cc:
Archival NDA: 20-508 Supplement
HFD-540/Division File
HFD-540/Dr. Wilkin
HFD-540/Dr. Okun
HFD-540/Dr. Cook
HFD-540/Mr. White
HFD-725/Dr. Huque
HFD-725/Dr. Al Osh
HFD-725/Mr. Thomson
HFD-340/Dr. Lepay
This review has 8 pages, including this signature page.
Chron.

APPEARS THIS WAY
ON ORIGINAL