CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

21-199

STATISTICAL REVIEW(S)

Statistical Review and Evaluation

NDA21-199
Name of Drug: Levofloxacin Ophthalmic Solution
Applicant: Santen Incorporated
Indication: Treatment of Bacterial Conjunctivitis
Documents Reviewed: Statistical Section (Vol.17-Vol.43) of NDA21-199
(Cover Letter Date 2/28/00)
Medical Reviewer: Wiley Chambers, M.D.
Reviewer: Laura Lu, Ph.D.
Date of Review: 7/27/2000

I. Background

NDA21-199 was submitted for approval of levofloxacin (LVFX) ophthalmic solution for treatment of bacterial conjunctivitis. Two Phase III clinical studies (Studies 03-003 and 03-004) were conducted to evaluate the clinical and microbial efficacy and safety of 0.5% LVFX versus 0.3% ofloxacin (OFLX) ophthalmic solution or placebo. Both studies were randomized, double-masked, multicenter studies which employed a 5-day dosing regimen. Additional studies conducted include a Phase I safety study, two Phase I pharmacokinetic studies, and a Phase II pilot study for 0.5% LFVX versus 0.3% OFLX. This review will concentrate on the two phase III studies.

II. Protocol

1. Study 03-003

This was a multicenter, randomized, active-controlled, double-masked comparison of 0.5% LVFX versus 0.3% OFLX in patients with bacterial conjunctivitis. The objective of this study was to evaluate the clinical and microbial efficacies, and safety of 0.5% LVFX ophthalmic solution compared to 0.3% OFLX ophthalmic solution for the treatment of bacterial conjunctivitis in adults and children one year of age and older.

Enrolled patients dosed with masked study medication for 5 days. Patients instilled one to two drops in the infected eye(s) every two hours, up to 8 times/day, while awake on Days 1 and 2, then one to two drops every four hours, up to 4 times/day, while awake on Days 3 through 5. The patients were examined at Visit 1 (Day 1/Baseline), Visit 2 (Day 3 or Day 4), and Visit 3 (Day 7±1).

The primary efficacy variables were clinical success and microbial success. Clinical outcome was rated on a four-point scale as resolved (0), improved (1), no change (2) or worse (3) based on change from baseline in cardinal signs at each follow-up visit. Microbial outcome was rated on a four-point scale as resolved (0), improved (1), no change (2) or worse (3) based on change from baseline in level of causative organisms at each follow-up visit. A clinical success was defined as a patient resolved or improved in clinical outcome at Visit 2 or Visit 3. A microbial success was defined as a patient resolved or improved in microbial outcome at Visit 2 or Visit 3.

The secondary efficacy variables were overall usefulness calculated from clinical outcome and microbial outcome as described in the table below.

Clinical Outcome	Microbial Outcome			
	Resolved (0)	Improved (1)	No Change (2)	Worse (3)
Resolved (0)	Extremely Useful (0)	Useful (3)	Minimally Useful (6)	Not Useful (9)
Improved (1)	Extremely Useful (1)	Useful (4)	Minimally Useful (7)	Not Useful (10)
No Change (2)	Useful (2)	Useful (5)	Minimally Useful (8)	Not Useful (11)
Worse (3)	Useful (3)	Minimally Useful (6)	Not Useful (9)	Not Useful (12)

The efficacy variables were analyzed by Cochran-Mantel-Haenszel statistic stratifying by centers. The primary patient population was the per-protocol population that includes only those patients whose Day 1 ocular specimen were culture-positive and had at least one post treatment evaluation. Last observation carried forward method was used in dealing with missing data.

Previous studies of LVFX (Phase II study) and OFLX (data available from Summary Basis of Approvals or equivalent documents provided under the Freedom of Information Act) provide the following estimates of the Day 7 microbial response rates for each treatment:

	Resolved	Improved	Worse/No Change
Ofloxacin	85%	15%	0%
Levofloxacin	60%	30%	10%

A sample size of 200 was planned (100 in each treatment group) to be able to detect the difference in estimated distribution of clinical outcome listed in the table above with power 90% and two sided α level 0.05 by Mantel-Haenszel Chi-square test. When a statistical significance was not found, it will be concluded that the difference between the two treatments was not greater than the meaningful difference on which the power calculations were based.

2. Study 03-004

The design of Study 03-004 was identical to that of Study 03-003 except for the controlled group, age requirement and sample size. Study 03-004 had two treatment groups: 0.5% LVFX and placebo. The age requirement in Study 03-004 was 2 years and older. Previous studies involving levofloxacin (Phase II study) and placebo (data available from Summary Basis of Approvals or equivalent documents provided under the Freedom of Information Act) provide the following estimates of the Day 3 microbial response rates for each treatment:

	Resolved	Improved	Worse/No Change
Levofloxacin	30%	60%	10%
Placebo	19%	30%	51%

A sample size of 96 was planned (48 in each treatment group) to be able to detect the difference in estimated distribution of clinical outcome listed in the table above with power 90% and two sided α level 0.05 by Mantel-Haenszel Chi-square test. When a statistical significance was not found, it will be concluded that the difference between the two treatments is not greater than the meaningful difference on which the power calculations were based.

III. Sponsor's Study Report

The results reported are for the per-protocol population. The efficacy result of the ITT population was consistent with that of the per-protocol population (see reviewer's comment #3 later).

1. Study 03-003

Patient Disposition

A total of 208 patients were included in the per-protocol population and only a few patients withdrew from the study. The detailed information is included in the table below.

Table 1. Patient Disposition in Study 03-003

	0.5% LVFX	0.3% OFLX
Enrolled	109	99
Completer	108(99.1%)	94(94.9%)
Dropouts	1(0.9%)	5(5.1%)
Reasons for Discontinuation		
Lack of Efficacy	1(0.9%)	1(1.0%)
Adverse Event	0(0.0%)	1(1.0%)
Other	0(0.0%)	3(3.1%)

Demographics

The patients were balanced in terms of age, sex and race between the two treatment groups. Detailed information for demographics are included in Table a.1 in Appendix A.

Primary Endpoints

The protocol specified that patients were to return for follow-up evaluations on Day 3 and Day 7. The protocols allowed for the Visit 2 follow-up evaluation to be done on Day 3 or Day 4. Similarly, the Visit 3 follow-up evaluation could be conducted on Day 6, 7 or 8. In practice, the logistical difficulties of scheduling follow-up visits around weekends and holidays resulted in data falling outside of the a priori visit windows defined by the protocols. To accommodate these data, the visit windows were expanded to study periods as described in the following table.

Table 2. Visit Windows

Day	By Visit	By Study Period
1	Baseline	Baseline
2	-	-
3*	Visit 2	Interim
4	Visit 2	Interim
5	-	Interim
6	Visit 3	Final
7*	Visit 3	Final
8	Visit 3	Final
9	-	Final
10	-	Final
Last observation (available postbaseline visit) carried forward	-	End of Study

* Best Day (the day defined in the protocol for each follow-up exam)

No statistical significance were found between the two treatment group in clinical success rate (p=0.584) at the end of study. Statistical significance was found in microbial success (p=0.051) in favor of 0.5% LVFX at the end of the study. The detailed results at interim period, final period, and the end of study are presented in the following table.

Table 3. Summary of Clinical and Microbial Success by Study Period

Study Period	Outcome	0.5% LVFX	0.3% OFLX	p-value
Interim (Day 3-5)	Clinical Success	62% (66/106)	68% (64/94)	0.364
	Microbial Success	91% (96/106)	91% (86/94)	0.646
Final (Day 6-10)	Clinical Success	93% (96/103)	95% (88/93)	0.977
	Microbial Success	92% (95/103)	85% (78/92)	0.054
End of Study (LOCF)	Clinical Success	92% (100/109)	91% (90/99)	0.584
	Microbial Success	93% (101/109)	85% (84/99)	0.051

Secondary Endpoints

A statistically significant difference in the distribution of overall usefulness scores was detected between the two treatment groups in favor of 0.5% LVFX (p=0.028) at the end of study, with 83% (90/109) of the patients in the 0.5% LVFX and 74% (73/99) of the patients in the 0.3% OFLX group receiving a score of extremely useful.

2. Study 03-004

Patient Disposition

A total of 117 patients were included in the per-protocol population and only a few patients withdrew from the study. The detailed information is included in the table below.

Table 4. Patient Disposition in Study 03-004

	0.5% LVFX	Placebo
Enrolled	60	57
Completer	59 (98.3%)	55 (96.5%)
Dropouts	1 (1.7%)	2 (3.5%)
Reasons for Discontinuation		
Lack of Efficacy	0(0.0%)	1(1.8%)
Adverse Event	0(0.0%)	0 (0.0%)
Other	1 (1.7%)	1 (1.8%)

Demographics

The patients were balanced in terms of age, sex and race between the two treatment groups. Detailed information for demographics are included in Table a.2 in Appendix A.

Primary Endpoints

Statistical significance was found in microbial success ($p<0.001$) in favor of 0.5% LVFX at the end of the study. No statistical significance was found in clinical success ($p=0.738$) at the end of the study. The detailed results are presented in the following table.

Table 5. Summary of Clinical and Microbial Success by Study Period

Study Period	Outcome	0.5% LVFX	Placebo	p-value
Interim (Day 3-5)	Clinical Success	63% (37/59)	47% (26/55)	0.155
	Microbial Success	97% (57/59)	51% (28/55)	<0.001
Final (Day 6-10)	Clinical Success	88% (52/59)	86% (48/56)	0.725
	Microbial Success	92% (54/59)	60% (33/55)	<0.001
End of Study (LOCF)	Clinical Success	87% (52/60)	84% (48/57)	0.738
	Microbial Success	92% (55/60)	60% (34/57)	<0.001

Secondary Endpoints

A statistically significant difference in the distribution of overall usefulness scores was detected between the two treatment groups in favor of 0.5% LVFX ($p<0.001$) at the end of study, with 82% (49/60) of the patients in the 0.5% LVFX and 47% (27/57) of the patients in the placebo group receiving a score of extremely useful.

IV. Reviewer's Comments

1. LVFX vs. Placebo

The primary efficacy variables specified in the protocol were clinical success rate (rate of clinical cure+improvement) and microbial success rate (rate of microbial eradication+improvement). In Study 03-004, no statistical significance was found between 0.5% LVFX and placebo in clinical success rate ($p=.738$) at the end of study. The sponsor also presented results for clinical cure rate in which a small p-value ($p=0.026$) was found between the two treatment groups. The results for clinical cure rate is presented in the table below.

Table 6. Summary of Clinical Cure by Study Period (Study 03-004)

Study Period	Clinical Cure Rate		P-value
	0.5% LVFX	Placebo	
Interim (Day 3-5)	27% (16/59)	24% (13/55)	0.642
Final (Day 6-10)	78% (46/59)	61% (34/56)	0.020
End of Study (LOCF)	77% (46/60)	60% (34/57)	0.026

Since cure rate was not a pre-specified primary efficacy variable, type I error rate inflation for making an efficacy claim based on this variable is a concern. This reviewer explored the magnitude of type I error rate inflation by the simulation procedure described in Appendix B. The simulation results show that LVFX claimed superiority vs. placebo at 5.3% of the time based on success rate and 9.2% of the time based on either success rate or cure rate, which means that if the treatment effect of the LVFX group was identical to that of the placebo group, the chance to claim superiority for LVFX over placebo by either clinical success rate or cure rate almost doubled the nominal type I error rate 0.05.

Another observation worth mentioning is that the cure rates of the two treatment groups at the interim period was numerically close (27% in LVFX and 24% in placebo).

2. LVFX vs. OFLX

In the study protocols, the sponsor stated that 'the lack of a statistically significant difference between two treatments in a study which had adequate power to detect a meaningful difference justifies a conclusion that the difference between the two treatments is not greater than the meaningful difference on which the power calculations were based'. This approach for equivalence claim is not consistent with that of the Agency's which requests the limits of the 95% confidence intervals for the treatment difference fall into a pre-specified equivalence

margin. Per this reviewer's request, the sponsor submitted the 95% confidence intervals for treatment difference at the end of study between LVFX and OFLX as listed in the table below.

Table 7. Confidence Interval (95%) for Difference in Clinical Success Rate and Cure Rate Between LVFX and OFLX at the End of Study

	0.5% LVFX N(%)	0.3% OFLX N(%)	Difference (LVFX-OFLX)* (%)	Lower Limit of 95% CI (%)	Upper Limit of 95% CI (%)
Cured	83 (76.15%)	75 (75.76%)	-1.55	-10.14	7.05
Success	100 (91.75%)	90(90.90)	0.73	-7.52	8.97

* results weighted by center

In order to claim equivalence between two drugs, there should be evidence that both drugs are superior to placebo and the effect sizes of the two drugs are close (within a pre-specified equivalence margin). Since LVFX did not show superiority to placebo in Study 03-004 and placebo arm was not included in Study 03-003 to provide essay sensitivity, the equivalence claim for LVFX and OFLX based on success rate is not valid.

3. ITT population vs. Per-Protocol Population

In Study 03-003 and 03-004, over half of the patients in ITT population were not included in the per-protocol population due to negative culture or no culture. The results in clinical outcome in ITT population were numerically similar to that in the per-protocol population as presented in the table below.

Table 8. Results of Clinical Outcome in ITT and Per-protocol Population

Endpoint	Study 03-003				Study 03-004			
	ITT		Per-Protocol		ITT		Per-Protocol	
	LVFX (N=206)	OFLX (N=206)	LVFX (N=109)	OFLX (N=99)	LVFX (N=121)	Placebo (N=117)	LVFX (N=60)	Placebo (N=57)
Clinical Cure	73% (151/206)	70% (144/206)	76% (83/109)	76% (75/99)	69% (83/121)	56% (65/117)	77% (46/60)	60% (34/57)
Clinical Success	89% (183/206)	88% (181/206)	92% (100/109)	91% (90/99)	80% (97/121)	83% (97/117)	87% (52/60)	84% (48/57)

V. Overall Conclusion

In study 03-004, the sponsor did not demonstrated superiority of 0.5% LVFX over placebo in clinical success rate. The sponsor provided results for the comparison in clinical cure rate between 0.5% LVFX and placebo (76% vs. 60%) in favor of LVFX with p-value=0.026. However, since cure rate was not a pre-specified primary efficacy variable, type I error rate inflation for making an efficacy claim based on this variable is a concern, and simulation confirms this concern.

In Study 03-003, equivalence between LVFX and OFLX based on success rate is not a valid claim since that LVFX did not show superiority over placebo. The clinical cure rates of LVFX and OFLX were numerically close (76.15% and 75.76%) with the 95% confidence interval for the difference being (-10.14%, 7.05%). Again, whether LVFX and OFLX are equivalent remains

questionable due to the lack of prespecified equivalence margins and essay sensitivity, and that the cure rate was not a prespecified endpoint.

/S/
Laura Lu, Ph.D.
Mathematical Statistician

Concur:

/S/ 7/27/00

Stan Lin, Ph.D.
Team Leader

CC:
NDA 21199
HFD-550/MO/Chambers
HFD-550/PM/Puglisi
HFD-550/Midthun
HFD-550/Div. File
HFD-725/Lu
HFD-725/Lin S.
HFD-725/Div. File

Appendix A. Tables

Table a.1 Patient Demographics in Study 03-003

	TREATMENT		
	0.5% LVFX	0.3% OFLX	COMBINED
Number of Patients:	109	99	208
AGE			
MEAN(SD)	29.28 (22.95)	28.24 (23.14)	28.78 (22.99)
MEDIAN	26.00	24.00	26.00
MIN-MAX	1-80	1-79	1-80
>16 years	67 (61.47)	62 (62.63)	129 (62.02)
12-16 years	7 (6.42)	6 (6.06)	13 (6.25)
2-11 years	30 (27.52)	26 (26.26)	56 (26.92)
<2 years	5 (4.59)	5 (5.05)	10 (4.81)
SEX: N(%)			
Female	66 (60.55)	61 (61.62)	127 (61.06)
Male	43 (39.45)	38 (38.38)	81 (38.94)
RACE: N(%)			
Caucasian	89 (81.65)	76 (76.77)	165 (79.33)
Non-Caucasian	20 (18.35)	23 (23.23)	43 (20.67)
Black	7 (6.42)	11 (11.11)	18 (8.65)
Asian	0 (0.00)	1 (1.01)	1 (0.48)
Hispanic	11 (10.09)	9 (9.09)	20 (9.62)
Other	2 (1.83)	2 (2.02)	4 (1.92)

Table a.2 Patient Demographics in Study 03-004

	TREATMENT		
	0.5% LVFX	PLACEBO	COMBINED
Number of Patients:	60	57	117
AGE			
MEAN(SD)	31.42 (22.26)	31.63 (22.95)	31.52 (22.50)
MEDIAN	29.50	29.00	29.00
MIN-MAX	2-91	2-76	2-91
>16 years	41 (68.33)	37 (64.91)	78 (66.67)
12-16 years	3 (5.00)	3 (5.26)	6 (5.13)
2-11 years	16 (26.67)	17 (29.82)	33 (28.21)
SEX: N(%)			
Female	38 (63.33)	25 (43.86)	63 (53.85)
Male	22 (36.67)	32 (56.14)	54 (46.15)
RACE: N(%)			
Caucasian	44 (73.33)	46 (80.70)	90 (76.92)
Non-Caucasian	16 (26.67)	11 (19.30)	27 (23.08)
Black	10 (16.67)	6 (10.53)	16 (13.68)
Asian	0 (0.00)	1 (1.75)	1 (0.85)
Hispanic	5 (8.33)	3 (5.26)	8 (6.84)
Other	1 (1.67)	1 (1.75)	2 (1.71)

Appendix B. Simulation Procedure in Assessing Type I Error Inflation

1). Generate random clinical outcome data for LVFX and placebo groups based on the observed distribution of clinical outcome in the placebo group: 60% resolved, 25% improved, and 15% unchanged or worse with n=60 (the planned sample size in the protocol) for each group.

2). Compare the clinical success rate and cure rate between the two treatment groups by T-tests.

3). Repeat 1) and 2) 2000 times and calculate the percentage of times that LVFX showed superiority over placebo in success rate at level 0.05, and also calculate the percentage of times that LVFX showed superiority over placebo in success rate and cure rate at level 0.05.