U.S. 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

CLINICAL STUDIES

NDA/Serial Number:	N22-159
Drug Name:	OraVerse 0.4mg/1.7mL cartridge
Indication(s):	Reversal of soft tissue anesthesia and associated functional deficits resulting from an intraoral submucosal injection of a local anesthetic containing a vasoconstrictor in [redacted] years patients
Applicant:	Novalar Pharmaceuticals, Inc
Date(s):	Received 04/09/07; User Fee 02/09/08
Review Priority:	Standard
Biometrics Division:	Division of Biometrics II/Office of Biostatistics
Statistical Reviewer:	Feng Zhou, M.S. (Statistical Reviewer)
Concurring Reviewers:	Dionne L. Price, Ph.D. (Statistical Team Leader)
Statistical Supervisor:	Thomas J. Permutt, Ph.D. (Statistical Division Director)
Medical Division:	Division of Anesthesia, Analgesia, and Rheumatology Products
Clinical Team:	Arthur Simone, M.D. (Medical Reviewer) Rigoberto Roca, M.D. (Medical Team Leader) Bob Rappaport, M.D. (Medical Division Dirctor)
Project Manager:	Geri Smith

Keywords: NDA review, clinical Studies

Table of Contents

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Novalar Pharmaceuticals has proposed OraVerse® (phentolamine mesylate) injection for "the reversal of soft tissue anesthesia and the associated functional deficits resulting from an intraoral submucosal injection of a local anesthetic containing a vasoconstrictor." The sponsor conducted seven studies including three phase 3 studies, referenced in the label, to evaluate the efficacy of NV-101. The sponsor claims that in all three studies, the data demonstrate the statistical and clinical superiority of phentolamine mesylate to sham for the efficacy parameters measured (time to recovery of normal lip sensation, time to a score of zero on the Soft Tissue Anesthesia Recovery Questionnaire (STAR-7), time to normal function, and time to recovery of normal tongue sensation).

Based on my collective evaluation of the NDA submission, I conclude that there is evidence of the effect of phentolamine mesylate in reversal of soft tissue anesthesia (STA) in subjects undergoing dental procedures involving the mandible and maxilla with an anesthetic/vasoconstrictor combination. The studies demonstrated a faster recovery of normal lip sensation, as well as normal abilities to smile, speak, drink refrain from drooling.

1.2 Brief Overview of Clinical Studies

The clinical development plan was introduced to the Division of Anesthesia, Analgesia, and Rheumatology Products by Novalar Pharmaceuticals via IND 65,095 and discussed during several meetings. Discussions during the meetings focused on the adequacy of the proposed primary endpoint, the choice of anesthetics, frequency of assessments, and the evaluation of the functional and mental impact of soft tissue anesthesia recovery.

The sponsor submitted this application on April 9, 2007 (NDA 22-159) in support of the proposed indications for the NV-101 [redacted], 0.4mg, [redacted] dosage strengths.

Table 1 summarizes the design and statistical results for the primary efficacy endpoint for the three studies used by the sponsor to support efficacy.

Table 1. Design and Statistical Results of Three Studies

Study / Center / Study Period	*Gender Mean Age (Range)*	*Study Design*	*No. of subjects by treatment group entered/completed*	*Primary Endpoints*	*Time Difference * 95% CI p-value*
NOVA 04-100 18 centers in U.S. 2/10/06–5/26/06	Male 120 Female 124 36 (12 – 92)	Randomized Blinded Controlled	NV-101 0.4mg: 89/89 NV-101 0.8mg: 33/33 NV-101 0 (sham): 122/122	Time to recovery of normal sensation of the lower lip	Sham vs. NV-101: Δ=85 in median (min) Hazard Ratio=3.21, 95%CI: (2.46, 4.20), p<0.0001
NOVA 04-200 16 centers in U.S. 2/10/06-6/2/06	Male 111 Female 129 38 (13 – 81)	Randomized Blinded Controlled	NV-101 0.4mg: 113/113 NV-101 0.8mg: 7/7 NV-101 0 (sham): 120/120	Time to recovery of normal sensation of the upper lip	Shem vs. NV-101: Δ = 82 in median (min) Hazard Ratio=3.05, 95%CI: (2.32, 4.0), p<0.0001
NOVA 05-PEDS 11 centers in U.S. 3/3/06-6/24/06	Male 75 Female 77 7.7 (4 – 11)	Randomized Blinded Controlled	NV-101 0.2mg: 74/74 NV-100 0.4mg: 22/22 NV-101 0 (sham): 56/56	Acceleration of the time to normal lip sensation in maxillary procedures	Shem vs. NV-101: Δ = 75 in median (min) Hazard Ratio=4.16, 95%CI: (2.53, 6.83), p<0.0001

* Results from reviewer's analysis; code: km_anal.sas.

1.3 Statistical Issues and Findings

There was one issue that arose during the review. About 50% of subjects had study procedure deviations related to the collection of Functional Assessments Battery data (a secondary efficacy variable). The sponsor reported the analysis results based on imputed data. I checked the protocol deviations case listings. Most deviations appeared to result from assessments not performed at the scheduled time point or missing assessments at some time points due to various reasons. I performed an additional analysis including only the subjects who did not have a study procedure deviation of the Functional Assessments Battery. The results are consistent with the sponsor's results and support the efficacy of NV-101.

My evaluation of the data supports the sponsor's conclusion of Studies NOVA 04-100 and NOVA 04-200. I am also in agreement with the sponsor's results in study NOVA PED-05. These studies provide evidence of the efficacy of phentolamine mesylate.

2. INTRODUCTION

2.1 Overview

NV-101 (phentolamine mesylate) is proposed "for the reversal of soft tissue anesthesia (STA) caused by local anesthetics containing a vasoconstrictor using intraoral injection techniques." The clinical development plan was introduced to the Division of Anesthesia, Analgesia, and Rheumatology Products by Novalar Pharmaceuticals via IND 65,095 and discussed during an End-of Phase 2 meeting (October 30, 2003), a Type A meeting (November 18, 2004), and a teleconference pertaining to the Special Protocol Assessment (May 4, 2005). Discussions during the meetings focused on the adequacy of the proposed primary endpoint, the choice of anesthetics, and the evaluation of the functional and mental impact of soft tissue anesthesia recovery. An additional focus of the latter meeting was the frequency of assessments. On September 13, 2005, the sponsor submitted the revised protocols (Protocol NOVA 04-100 and Protocol NOVA 04-200) incorporating changes recommended during the numerous correspondences and meetings between the sponsor and the Division. The proposed statistical analysis plans for both studies were acceptable.

The sponsor submitted this application on April 9, 2007 (NDA 22-159) in support of the proposed indications for the NV-101 [redacted], 0.4mg, and [redacted] dosage strengths.

The sponsor's submission included seven studies. I will focus on three phase 3 studies which are outlined in Table 2.

Table 2. Clinical Trials

Study/Center/ Study Period	***Study Design***	***Key Inclusion Criteria***	***No. of subjects by treatment group entered/completed***	***Primary Endpoints***
NOVA 04-100 18 centers in U.S. 2/10/06–5/26/06	Randomized Blinded Controlled	Subjects (mandible) aged 12 years and older undergoing standard dental procedures	NV-101 0.4mg: 89/89 NV-101 0.8mg: 33/33 NV-101 0 (sham): 122/122	Time to recovery of normal sensation of the lower lip
NOVA 04-200 16 centers in U.S. 2/10/06-6/2/06	Randomized Blinded Controlled	Subjects (maxilla) aged 12 years and older undergoing standard dental procedures	NV-101 0.4mg: 113/113 NV-101 0.8mg: 7/7 NV-101 0 (sham): 120/120	Time to recovery of normal sensation of the upper lip
NOVA 05-PEDS 11 centers in U.S. 3/3/06-6/24/06	Randomized Blinded Controlled	Pediatric subjects aged 4 to 11 years undergoing standard dental procedures	NV-101 0.2mg: 74/74 NV-100 0.4mg: 22/22 NV-101 0 (sham): 56/56	Acceleration of the time to normal lip sensation in maxillary procedures; Acceleration of time to normal tongue sensation in mandibular procedures

* Results from reviewer's analysis; code: demo.sas.

2.2 Data Sources

Documents reviewed were accessed from the CDER document room at: \\...\N22159\

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

The main body of my evaluation of efficacy will encompass both adult studies (NOVA 04-100 and NOVA 04-200) and discuss one pediatric study (NOVA 05-PEDS).

3.1.1 Study NOVA 04-100 and NOVA 04-200

Study Design and Endpoints

Both NOVA 04-100 and NOVA 04-200 were randomized, multi-center, double-blind, controlled studies evaluating the safety and effectiveness of NV-101 in patients aged 12 years and older who underwent dental procedures. The studies were nearly identical with the exception of the permissible dental procedures. In Study NOVA 04-100, eligible patients underwent mandibular procedures while Study NOVA 04-200 enrolled patients undergoing maxillary procedures.

Initially, potential study participants were randomized to a local anesthetic consisting of lidocaine 2% with epinephrine, articaine 4% with epinephrine, prilocaine 4% with epinephrine, or mepivacaine 2% with levonordefrin. According to Novalar, lidocaine 2% with epinephrine is the most commonly used anesthetic in dental practice; therefore, the sponsor randomized to this anesthetic or another anesthetic via a 2:1 allocation ratio. Following completion of the dental procedure (either maxillary or mandibular depending on the study), eligible patients additionally were randomized to receive NV-101 or a sham injection in a 1:1 allocation ratio in the location at which the anesthetic was administered. The number of treatment cartridges (one or two) corresponded to the number of cartridges of local anesthetic administered. Randomization to study treatment was stratified according to study center, age (12-17, 18-64, and 65+), anesthetic (lidocaine, other), and the number of cartridges of anesthetic administered (1 or 2). After study drug administration, patients remained at the study site for 5 hours. During this period, patients assessed lip numbness every 5 minutes via a standardized palpation procedure. Moreover, patients completed the Soft Tissue Anesthesia Recovery (STAR-7) questionnaire prior to administration of study drug and every 30 minutes during the 5-hour period. The instrument was designed to evaluate the functional and mental impact of soft tissue anesthesia. Specifically, the questionnaire consisted of seven items related to eating, drinking, speaking, smiling, and drooling. A total score of zero on the questionnaire denoted normal sensation. Patients also completed the Functional Assessments Battery (FAB) which evaluated functional impairment via patient and observer assessments of speaking, smiling, drooling, and drinking. Initially, the former three components were tested every 5 minutes. Once these components were rated as normal by both the patient and the observer, drinking was subsequently added to the 5-minute evaluations. A telephone follow-up was conducted within 2 days of study drug administration.

The primary efficacy endpoint was the time to recovery of normal lip (either lower lip or upper lip depending on the study) sensation. The time was measured from administration of study drug

to the first of two consecutive reports of normal sensation in the lip (measured by the standardized palpation procedure). The recovery of normal lip sensation also was considered to have occurred if the lip sensation test was rated normal at the subject's final evaluation and the rating from the preceding assessment was other than normal (i.e., not done, numb, or tingling). The primary endpoint was censored for participants that had not experienced a return to normal sensation by the end of the 5-hour period.

Secondary variables included the time to a score of zero on the STAR-7 questionnaire and time to normal function (as measured via the FAB). The STAR-7 score was calculated by adding the responses pertaining to items 2 (uncomfortable), 3 (biting), 4 (drinking), 6 (speaking), 7 (smiling), 8 (drooling), and 11 (appearance to others) on the STAR questionnaire. Similar to the primary efficacy variable, the time to STAR-7 score of zero was calculated by the number of minutes elapsed from the administration of study drug to the first of 2 consecutive STAR-7 scores of zero. This event was also considered to have occurred if the subject's last reported STAR-7 score was zero and the score from the preceding assessment was greater than zero or missing. Subjects who did not meet these criteria before the end of the 5-hour observation period were censored at the last time the subject completed the STAR questionnaire. The time to return of normal function (measured via the FAB) was calculated by the number of minutes elapsed from the administration of study drug to the first of 2 consecutive assessments in which both the subject and the observer rated smiling, speaking, and drinking as normal and drooling was not present. The sponsor stated, "The return of normal function was also considered to occur if all functional tests were rated normal or not present for the subject's last functional assessment battery and 1 or more of these tests from the preceding assessment was rated other than normal (i.e., not done or abnormal). Subjects who did not meet these criteria before the end of the 5-hour observation period were right-censored at the last time the subject completed the FAB with none of the individual subject or observer rated assessments missing."

Statistical Methodologies

The primary analysis employed a log rank statistic stratified by the local anesthetic and the number of cartridges of anesthetic administered. A proportional hazards model was used to test for the presence of treatment by strata interactions. The difference in the time to recovery of normal sensation between treatment groups was examined using Kaplan-Meier estimates. The sponsor investigated the homogeneity of the results among centers and age groups in separate analyses. The sponsor stated, "Although the randomization is also stratified by age group and study center, the primary analysis would not include these variables as stratification factors since it is anticipated that the number of subjects in one or more stratum within a give study center would be less than 2, thus resulting the efficiency loss due to deleted cases in the calculation of the stratified test statistic."

A secondary analysis employed a Weibull accelerated failure time model. This model allows determination of relative increases or decreases in the time to an event (termed the event time ratio).

A sequential step-down procedure was employed for inferential testing of the endpoints. If the primary endpoint reached statistical significance, the sequential procedure would lead to testing the hypothesis on the first secondary endpoint (STAR-7). If the STAR-7 endpoint reached statistical significance, the sequential procedure would lead to testing the hypothesis on the FAB. An additional secondary endpoint in Study NOVA 04-100 was the time to recovery of normal tongue sensation. The endpoint was measured in the same manner as other variables and was included in the step-down procedure having the least rank in order of importance.

According to the sponsor, "The LOCF method was used to impute missing item/component scores for the STAR-7 and FAB assessments. After the LOCF approach was used, in some cases, a single item score for STAR-7 or FAB was still missing. If the number of missing item scores was less than or equal to 20% of total number of items, the item score was imputed as the average of nonmissing component scores at each respective timepoint. If the number of missing item score was greater than 20%, the item scores were not imputed and therefore, the total score would be missing."

The primary analysis was conducted on the intent-to-treat (ITT) population including all randomized patients. The secondary analyses were conducted on the following modified intent-to-treat (MITT) analysis sets (Table 3):

- The MITT-STAR analysis set included all randomized subjects who had a STAR-7 score greater than zero for the STAR questionnaire given immediately before the randomization of study drug.
- The MITT-FAB analysis set included all randomized subjects who were rated abnormal by both the subject and the observer for at least one of the individual functional test (not necessarily the same test) given immediately before the randomization of study drug.
- The MITT-Tongue Sensation analysis set included all randomized subjects who had numbness of the tongue based on the tongue sensation test performed immediately before the randomization of study drug.

Based on previous studies, the sponsor determined that a sample of size 240 would be required to detect a 35% (or more) reduction in median recovery time with 90% power. The sample size was derived with the anticipation that 6% (or less) of the participants might not achieve a return to normal sensation in the lip by the end of the 5-hour period.

Patient Disposition, Demographic and Baseline Characteristics

As shown in the Table 3, more than 50% of patients were found to have protocol deviations. In nearly all patients with deviation, the deviations were related to study procedures. Of the 422 procedural deviations, 220 (53%) involved use of the FAB tool. The sponsor attributed the study procedure deviations to the complexity of the FAB data collection schedule. The sponsor stated, "These study procedure deviations were minor in scope and would not have affected the overall conduct of the study or the integrity of the data. In particular, the deviations that occurred in the collection of the FAB data did not change the overall interpretation of the FAB results, as the

type and effect of the deviations were balanced between the 2 treatment groups." I checked the protocol deviations case listings. Most deviations appeared to result from assessments not performed at the scheduled time point or missing assessments at some time points due to various reasons. I considered the sponsor's explanation and conclusion regarding the impact of the deviations to be acceptable.

Table 3. Patients' Accountability N (%)

	NOVA 04-100		NOVA 04-200	
	NV-101	Sham	NV-101	Sham
Randomized patients	122	122	120	120
Completed treatment period	122	122	120	120
Discontinued	0	0	0	0
Analysis Population				
ITT	122	122	120	120
MITT – STAR	118 (96.7%)	121 (99.2%)	109 (90.8%)	111 (92.5%)
MITT – FAB	103 (84.4%)	103 (84.4%)	100 (83.3%)	89 (74.2%)
MITT – Tongue	93 (76.2%)	103 (84.4%)	–	–
Safety	122	122	120	120
Number of ITT patients with a protocol deviation	69 (56.6)	66 (54.1)	67 (55.8)	69 (57.5)
Inclusion/Exclusion criteria	0	1	0	1
Study drug	6	4	0	1
Randomization	0	0	0	0
Study procedure	69 (56.6)	65 (53.3)	66 (55.0)	68 (56.7)
FAB	58	51	53	56
Others	11	14	13	12
Blinding	2 (1.6)	3 (2.5)	1 (0.8)	3 (2.5)

* Results from reviewer's analysis; code: demo.sas.

Note: patients could have more than 1 type of deviation; patients with more than 1 deviation in a category were counted once in that category.

Of the 494 patients initially enrolled in both studies, 484 patients were subsequently randomized to study drug and treatment with NV-101 or sham. All subjects completed the study. Table 4 summarizes demographics and baseline characteristics for the ITT population. The ages of patients ranged from 12 to 92 with a mean age of 38. In both studies, 78% of patients were Caucasian, and 11% were African-American. Forty-eight percent of the population was male. The demographic and baseline characteristics were similar across the treatment groups within each study with a few exceptions. There was a 10% difference in gender between the treatment groups in Study NOVA 04-100, and the mean age was 3 years younger in the NV-101 treatment group than sham group. Only 5% of patients required 2 cartridges (injection) of anesthetic in Study NOVA 04-200 while 25% of patients required 2 cartridges (injection) of anesthetic in Study NOVA 04-100. Most subjects received the primary injection of anesthetic by inferior alveolar nerve block in Study NOVA 04-100 while most of the subjects received the primary anesthetic by mental-incisive block in Study NOVA 04-200.

Table 4. ITT Subjects' Demographics and Baseline Characteristics by Treatment

	NOVA 04-100		NOVA 04-200	
	NV-101 (n=122)	Sham (n=122)	NV-101 (n=120)	Sham (n=120)
Sex				
Male	66 (54.1%)	54 (44.3%)	56 (46.7%)	55 (45.8%)
Female	56 (45.9%)	68 (55.7%)	64 (53.3%)	65 (54.2%)
Race Group				
White	99 (81.1%)	96 (78.7%)	92 (76.7%)	90 (75.0%)
Black	11 (9.0%)	12 (9.8%)	15 (12.5%)	17 (14.2%)
Asian	1 (0.8%)	8 (6.6%)	1 (0.8%)	5 (4.2%)
Native Hawaiian/Pacific Islander	0	0	1 (0.8%)	2 (1.7%)
American Indian	3 (2.5%)	1 (0.8%)	0	1 (0.8%)
Other	8 (6.6%)	5 (4.1%)	11 (9.2%)	5 (4.2%)
Age Group				
12 – 17	16 (13.1%)	15 (12.3%)	10 (8.3%)	14 (11.7%)
18 – 64	93 (76.2%)	93 (76.2%)	94 (78.3%)	94 (78.3%)
65+	13 (10.7%)	14 (11.5%)	16 (13.3%)	12 (10.0%)
Mean (SD)	35.2 (18.3)	38.1 (18.8)	38.5 (18.5)	38.1 (18.0)
Median	29.5	33.0	34.5	33.5
Range	12 – 92	13 - 84	13 – 81	12 – 78
Anesthetic Administrated				
Lidocaine/Epinephrine	82 (67.2%)	81 (66.4%)	79 (65.8%)	80 (66.7%)
Articaine/Epinephrine	16 (13.1%)	12 (9.8%)	17 (14.2%)	10 (8.3%)
Prilocaine/Epinephrine	13 (10.7%)	14 (11.5%)	14 (11.7%)	13 (10.8%)
Mepivacaine/Levonordefrin	11 (9.0%)	15 (12.3%)	10 (8.3%)	17 (14.2%)
Number of Anesthetic Injection(s)				
1	91 (74.6%)	91 (74.6)	113 (94.2%)	116 (96.7%)
2	31 (25.4%)	31 (25.4%)	7 (5.8%)	4 (3.3%)
Dental Procedure				
Cavity	88 (72.1%)	79 (64.8%)	78 (65.0%)	88 (73.3%)
Crown	0	3 (2.4%)	3 (2.5%)	4 (3.3%)
Periodontal maintenance	34 (27.9%)	40 (32.8%)	39 (32.5%)	28 (23.3%)
Mouth Quadrant				
Right lower (upper for 04-200)	54 (44.3%)	59 (48.4%)	63 (52.5%)	67 (55.8%)
Left lower (upper for 04-200)	68 (55.7%)	63 (51.6%)	57 (47.5%)	53 (44.2%)
Primary Injection Type				
Inferior alveolar nerve block	96 (78.7%)	98 (80.3%)	6 (5.0%)	2 (1.7%)
Mental-incisive block	21 (17.2%)	24 (19.7%)	109 (91%)	111 (93%)
Superior anterior alveolar nerve block	0	0	9 (7.5%)	17 (14.2%)
Supraperiosteal injection	5 (4.1%)	0	105 (87.5%)	101 (84.2%)

* Results from reviewer's analysis; code: demo.sas.

Results and Conclusions

In Study NOVA 04-100, 243/244 subjects reported numbness in the lower lip after injection of the local anesthetic and completion of the dental procedure but prior to injection of the randomized study drug. The subject (100-05-014) who did not report numbness was randomized to sham and reported tingling in the lower lip after injection of 2 cartridges of local anesthetic/vasoconstrictor (lidocaine/epinephrine).

The median elapsed time between injections of anesthetic and study drug was 47 minutes (range: 13 to 83 minutes) for the overall cohort in two studies (Table 5). I additionally created an

indicator variable (Elaps_g45) for patients who had more than 45 minutes elapsed time and included this variable in the Cox proportional hazards model.

Table 5. Elapsed Time

	NOVA 04-100		NOVA 04-200	
	NV-101 (n=122)	Sham (n=121)	NV-101 (n=120)	Sham (n=119)
Median elapsed time between injections of anesthetic and study drug	44	52	45	47
Range	(17, 78)	(20, 74)	(14, 81)	(13, 83)
# of Patients whose Elapsed Time > 45 min (%)	59 (48%)	74 (61%)	58 (48%)	65 (54%)

* Results from reviewer's analysis; code: km_anal.sas

In studies NOVA 04-100 and NOVA 04-200, the Kaplan-Meier estimated median time to recovery of normal sensation was 70 minutes and 50 minutes for subjects randomized to sham in the respective studies and 155 minutes and 130 minutes for subjects randomized to NV-101, respectively. The differences between these times were significant ($p<0.001$) using a log-rank test stratified for the number of cartridges and type of anesthetic/vasoconstrictor. The effect of NV-101 represented a reduction of 55% in median time to recovery of normal lower lip sensation for subjects treated with NV-101 compared with placebo. Similarly, subjects randomized to NV-101experienced a greater reduction (62%) in the median time to recovery of normal upper lip sensation compared to subjects randomized to placebo. Figure 1 and Figure 2 present the Kaplan-Meier plots comparing the time to recovery of normal sensation among the treatment groups. The y-axis corresponds to the proportion of patients without normal sensation.

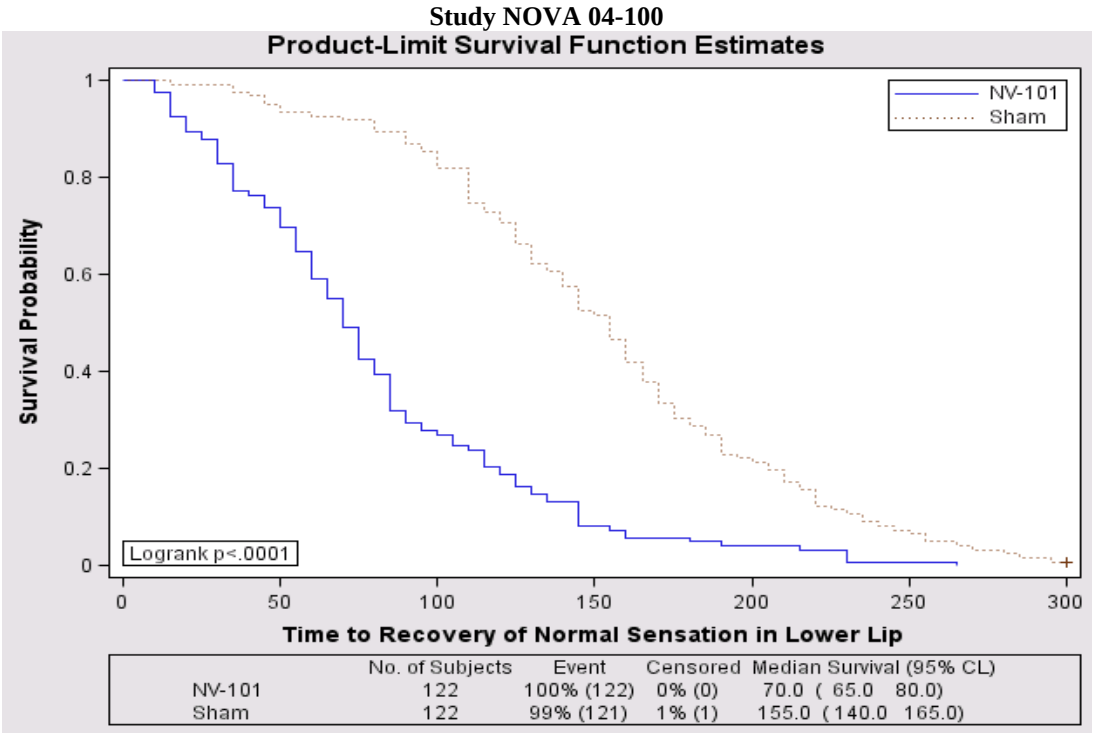
Table 6. Log-Rank Analysis of Time to Recovery of Normal Sensation in Lower (or Upper) Lip

	NOVA 04-100		NOVA 04-200	
	NV-101 (n=122)	Sham (n=121)	NV-101 (n=120)	Sham (n=119)
Median Time to recovery of normal sensation of lower (upper) lip[1] (min.)	70	155	50	132.5
95% CI	(65, 80)	(140, 165)	(45, 60)	(115, 145)
NV-101 vs. placebo				
Hazard ratio[2]		3.26		3.0
95% CI		(2.47, 4.29)		(2.3, 4.0)
p-value		<0.001		<0.001

Note: 1. Kaplan-Meier estimated; 2. Log-Rank test stratified for the number of cartridges and type of anesthetic/vasoconstrictor.

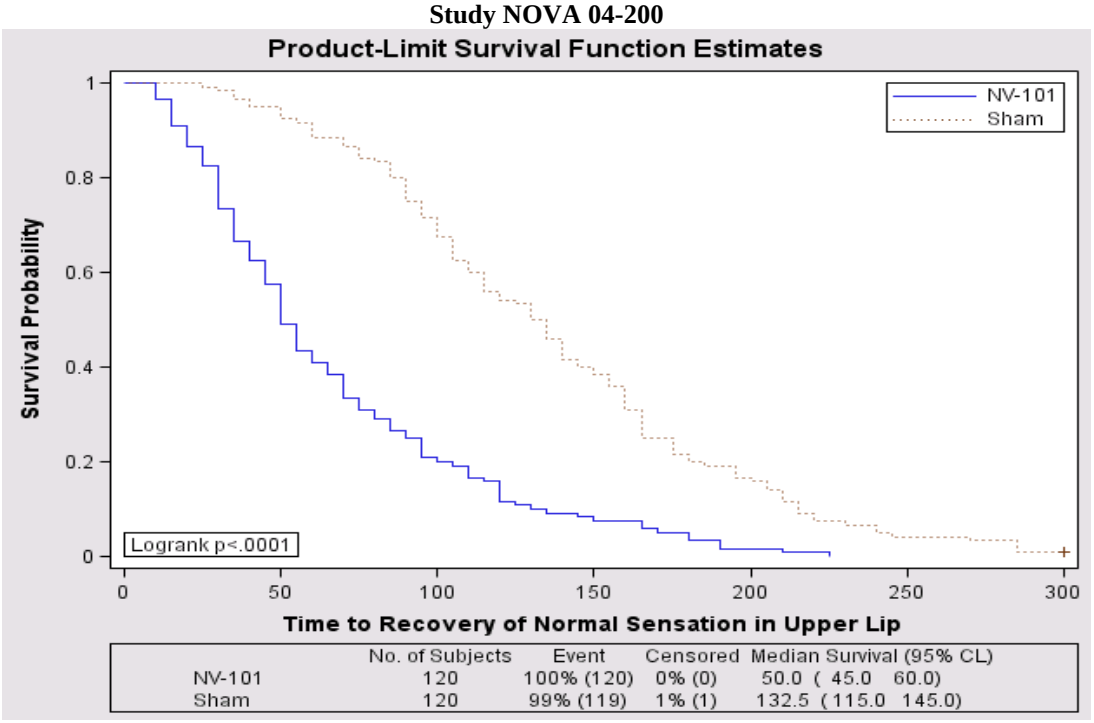
* Results from reviewer's analysis; code: km_anal.sas.

Figure 1. Kaplan-Meier Plot of Time to Recovery of Normal Sensation in Lower Lip (ITT)



* Results from reviewer's analysis; code: km_anal.sas.

Figure 2. Kaplan-Meier Plot of Time to Recovery of Normal Sensation in Upper Lip (ITT)



* Results from reviewer's analysis; code: km_anal.sas.

The Cox proportional hazards model was employed to evaluate the effects of treatment, anesthetic, and number of cartridges on the time to return to normal sensation. The hazard ratio between the 2 treatment groups was also computed and indicated the likelihood of achieving normal sensation in the lower or upper lip. The model proposed by the sponsor included fixed effects of treatment group, anesthetic (lidocaine, all others pooled), and the number of anesthetic cartridges administered (1 or 2). The homogeneity of the treatment effect across strata was evaluated by testing for the presence of treatment group by strata interaction. The interaction effects were not statistically significant; therefore, the interaction terms were not included in the final model. I additionally created an indicator variable (Elaps_g45) for patients who had more than 45 minutes elapsed time and included this variable in the proportional hazards model.

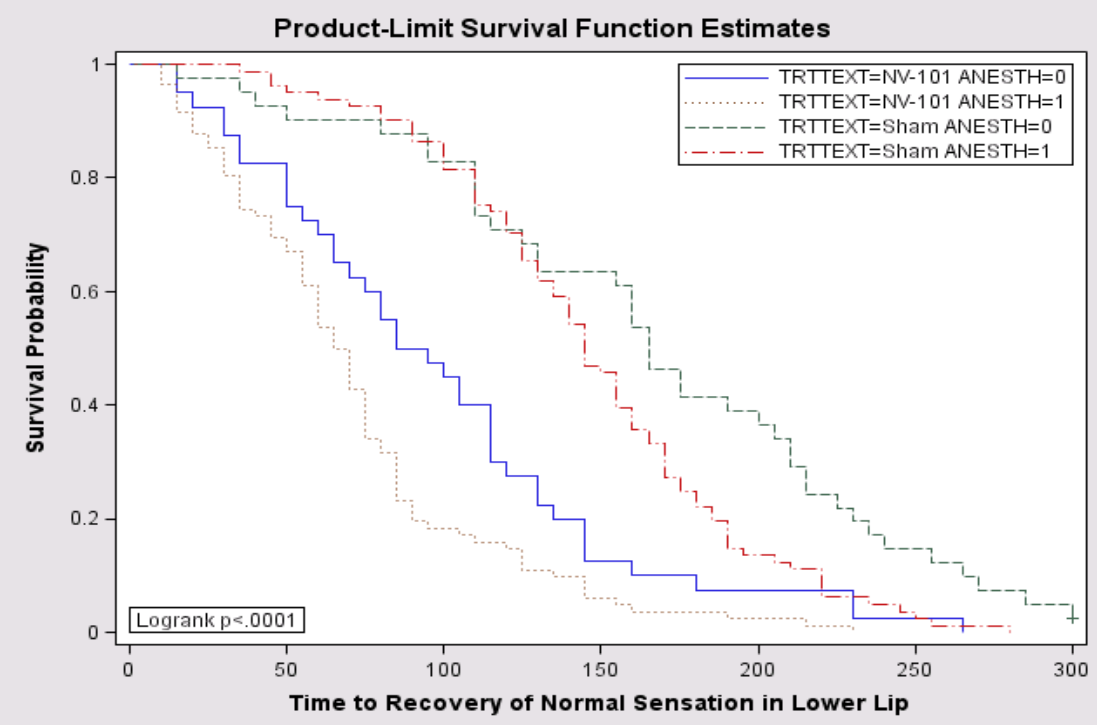
The results of the Cox proportional hazards model (including a term for elapsed time) are shown in Table 7. The hazard ratio for the treatment group was 3.28 implying that after adjustment for type of local anesthetic and number of anesthetic cartridges, subjects in the NV-101 treatment group were 3 times as likely to achieve normal sensation during the 5-hour observation period as subjects treated with sham. The model also predicted a hazard ratio of 1.68 for the effect of the type of local anesthetic (lidocaine vs. other anesthetic). This effect was statistically significant in study NOVA 04-100, suggesting that subjects who were anesthetized with the combination of lidocaine/epinephrine were 1.7 times more likely to achieve normal sensation than subjects who were anesthetized with the other 3 anesthetics combined, regardless of treatment group. This result was not evident in study NOVA 04-200 and is further illustrated by a Kaplan-Meier plot of the data, as shown in Figure 3 and Figure 4.

Table 7. Cox Proportional Hazards Model for Time to Recovery of Normal Sensation of the Lip

Variable	***NOVA 04-100***		***NOVA 04-200***	
	Hazard Ratio	**p-value 95%CI**	**Hazard Ratio**	**p-value 95%CI**
Treatment (NV-101 vs. sham)	3.28	<0.0001 (2.50, 4.30)	3.05	<0.0001 (2.32, 4.00)
Anesthetic (Lidocaine vs. Others)	1.68	0.0002 (1.27, 2.22)	1.06	0.677 (0.80, 1.40)
Number of Anesthetic Cartridges Adm. (1 vs. 2)	1.19	0.237 (0.89, 1.59)	1.43	0.256 (0.77, 2.65)
Elapsed time (min) (≤ 45 vs. > 45)	1.16	0.256 (0.90, 1.50)	1.02	0.879 (0.79, 1.32)

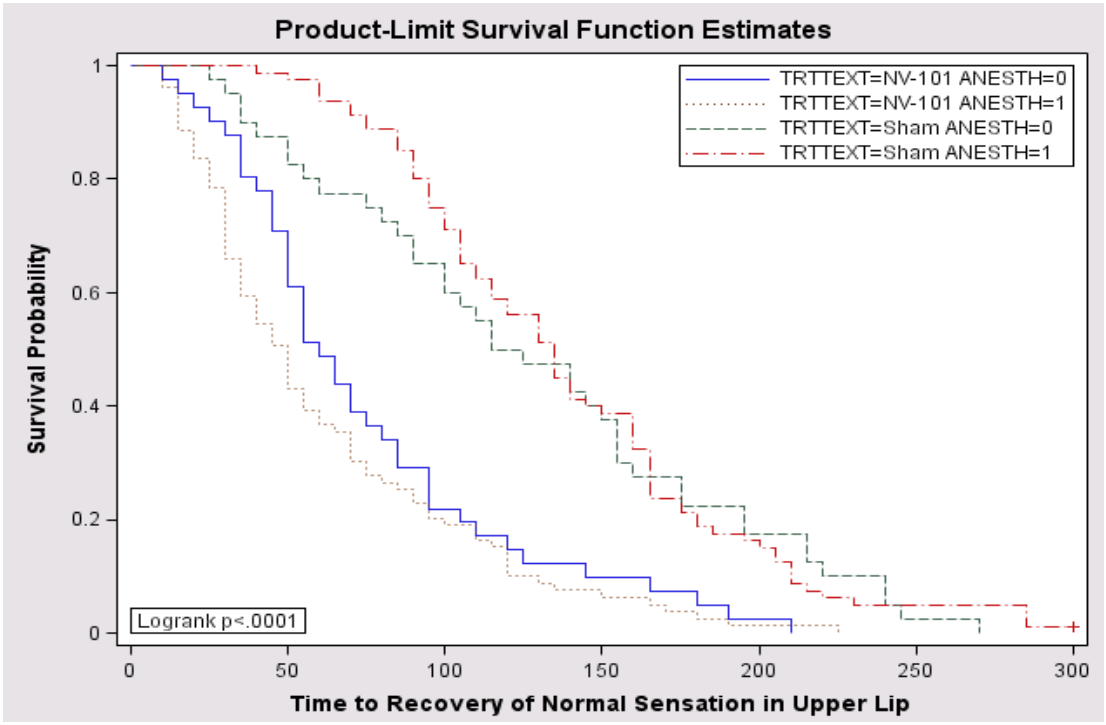
* Results from reviewer's analysis; code: km_anal.sas.

Figure 3. Kaplan-Meier Plot of Time to Recovery of Normal Sensation in Lower Lip



* Results from reviewer's analysis; code: km_anal.sas.

Figure 4. Kaplan-Meier Plot of Time to Recovery of Normal Sensation in Upper Lip



* Results from reviewer's analysis; code: km_anal.sas.

Secondary analysis – Weibull Accelerated Failure Time (AFT) model

A Weibull AFT model was used as an additional method for analysis of the primary endpoint. This model allows determination of relative increases or decreases in the time to an event (termed the event time ratio). Results of the model (Table 8) predicted an estimated time to event ratio for NV-101 versus sham of 0.57 (Study NOVA 04-100) and 0.53 (Study NOVA 04-200), indicating that NV-101 accelerated the time to normal sensation by 43% and 47% compared with sham. These results are consistent with the Cox proportional hazards model and the Kaplan Meier estimates.

Table 8. Weibull AFT Model for Time to Recovery of Normal Sensation of the Lip

Variable	NOVA 04-100			NOVA 04-200		
	e^{β}	SE	p-value	e^{β}	SE	p-value
Treatment (NV-101 vs. sham)	0.57	0.06	<.0001	0.53	0.07	<.0001
Anesthetic (Lidocaine vs. Others)	0.79	0.07	0.0004	0.97	0.08	0.6858
Number of Anesthetic Cartridges Adm. (1 vs. 2)	0.91	0.07	0.1964	0.82	0.17	0.2549

* Results from reviewer's analysis; code: wb_anal.sas.

Secondary variables -

Secondary variables included the time to STAR-7 score of zero and time to normal function (as measured via the FAB). Study NOVA04-100 also included the time to normal tongue sensation. As shown in Table 9, the secondary variables provided further evidence of the efficacy of NV-101 in reversing soft tissue anesthesia.

Table 9. Log-Rank Analysis of Secondary Efficacy Variables

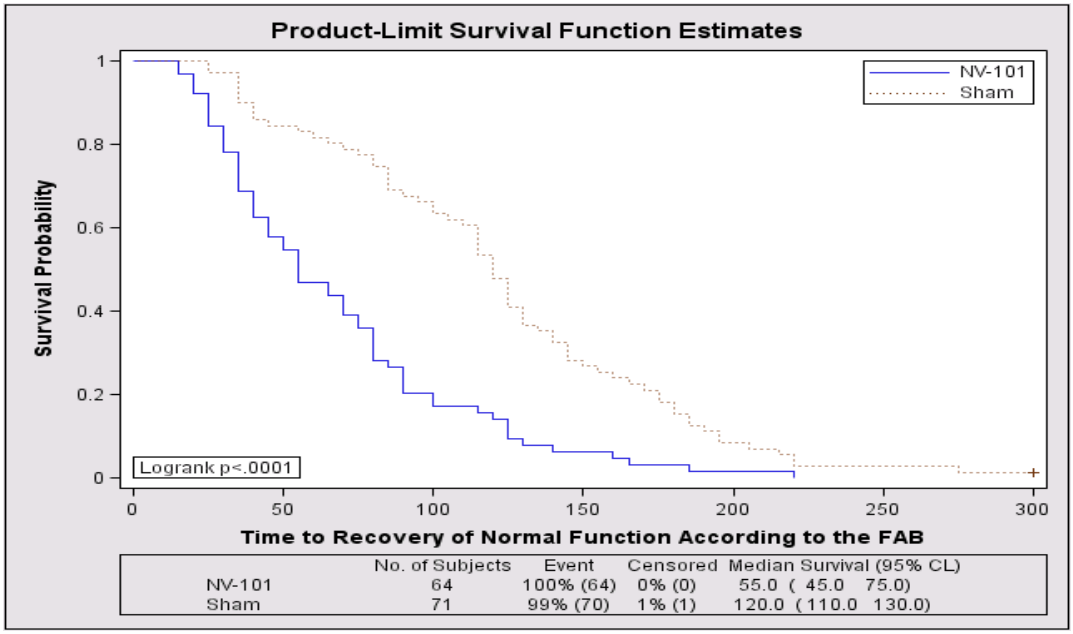
	Treatment				Treatment Comparison		
Time-to-Event (min)	NV-101		Sham		NV-101 vs. Sham		
Study NOVA 04-100	**N 122**	**Median[1] 95% CI**	**N 122**	**Median 95% CI**	**Time Difference**	**Reduction (%)**	**p-value[2]**
STAR-7 = zero	118	90 (60, 90)	121	150 (120, 150)	60	40.0%	<0.001
Normal FAB	103	65 (55, 75)	103	120 (110, 130)	55	45.8%	<0.001
Normal tongue sensation	93	60 (55, 70)	103	125 (110, 135)	65	52.0%	<0.001
Study NOVA 04-200	**N 120**	**Median 95% CI**	**N 120**	**Median 95% CI)**	**Time Difference**	**Reduction (%)**	**p-value**
STAR-7 = zero	109	60 (60, 90)	111	120 (120, 150)	60	50.0%	<0.001
Normal FAB	100	60 (55, 65)	89	110 (85, 130)	50	45.5%	<0.001

Note: 1. Kaplan-Meier estimated; 2. Log-Rank test stratified for the number of cartridges and type of anesthetic/vasoconstrictor. * Results from reviewer's analysis; code: km_anal.sas.

The review team expressed concern regarding the amount of missing data for FAB assessments. The sponsor submitted the imputed data set for STAR-7 and FAB assessments. I could not verify the imputation, but I confirmed the sponsor's analysis results using the submitted data set. I performed the analysis excluding subjects who had a FAB deviation based on the sponsor's list

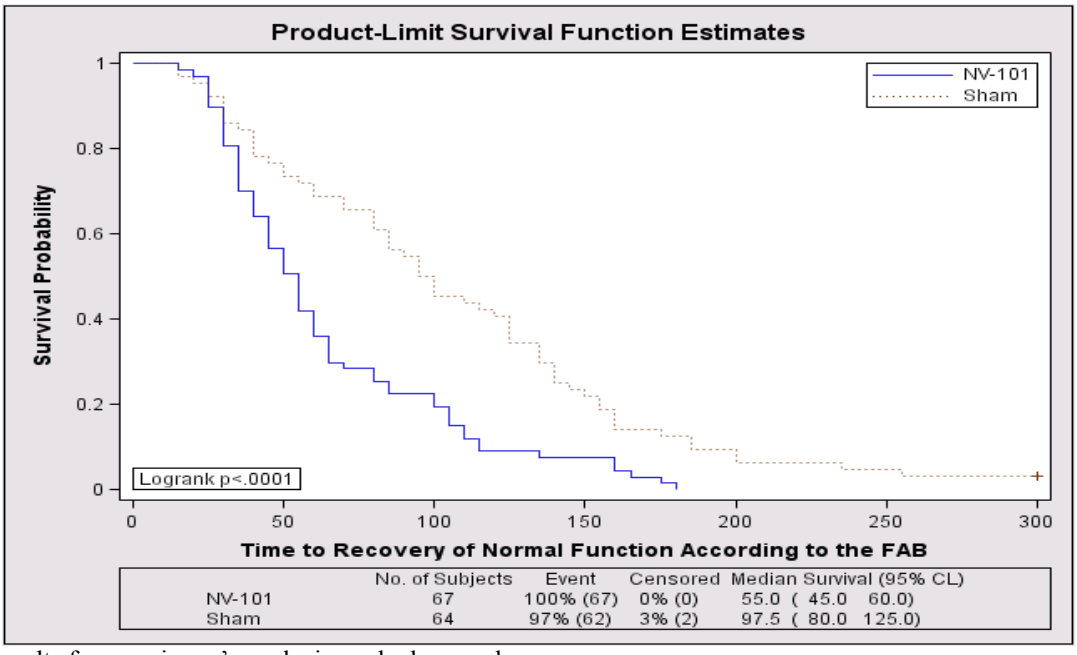
(16.2.2 Protocol Deviations). The results of the additional analysis including only subjects without a FAB deviation support the efficacy of NV101 in terms of the FAB endpoint. Figure 5 and Figure 6 display the Kaplan-Meier curves excluding imputed data. The results of the additional analysis are consistent with the results based on imputed data.

Figure 5. Kaplan-Meier Plot of Time to Recovery of Normal FAB (Study NOVA 04-100)



* Results from reviewer's analysis; code: km_anal.sas.

Figure 6. Kaplan-Meier Plot of Time to Recovery of Normal FAB (Study NOVA 04-200)



* Results from reviewer's analysis; code: km_anal.sas.

3.1.2 Study NOVA 05-PEDS

Study Design and Endpoints

NOVA 05-PEDS was a phase 2, multi-center, randomized, blinded, controlled study to evaluate the safety and efficacy of NV-101 used for reversal of STA in pediatric subjects (4 to 11 years old) undergoing dental or periodontal maintenance procedures in the mandible or maxilla under local anesthesia. Following completion of the dental procedure, eligible subjects were randomized to study drug at a 2:1 ratio (NV-101 or sham). Subjects weighing ≥ 15 kg and < 30 kg received a half cartridge of 2% lidocaine with 1:100,000 epinephrine and a half cartridge of NV-101 (0.2 mg phentolamine mesylate), and subjects weighing ≥30 kg received half or a whole cartridge of 2% lidocaine with 1:100,000 epinephrine and half or a whole cartridge of NV-101 (0.2 or 0.4 mg phentolamine mesylate, respectively). All subjects were contacted by telephone on Day 2 or Day 3 for follow-up of adverse events and concomitant medications.

The sponsor stated,

> The primary objective of this study was to evaluate the safety and tolerability of NV-101 in pediatric patients (4-11 years old) undergoing mandibular or maxillary dental procedures; thus, the sample size justification for this study was based on the probability of detecting potential adverse events that might have occurred in the NV-101 treatment group. If 100 subjects were enrolled in the NV-101 arm of the study, there would be a 95% confidence level of observing at least one occurrence of a specific adverse event if the true proportion of subjects that would develop this adverse event in the population was 3%. This study was not prospectively powered to detect treatment differences in the efficacy endpoints. The study was considered complete when approximately 150 subjects had been randomized to study drug (NV-101 or sham) and had completed the procedures of the protocol.

As secondary objectives for subjects 6 to 11 years of age who were trainable in standardized palpation procedures, the study determined if NV-101 accelerates the time to normal lip sensation (or normal tongue sensation in mandibular procedures) as measured by palpation at screening, before randomization to study drug, and every 15 minutes for 4 hours after completion of study drug administration, starting at 15 minutes after study drug administration.

Statistical Methodologies

> The time to recovery of normal sensation of the lip was summarized descriptively by treatment group using the Kaplan-Meier method. The estimated median for each treatment group and the corresponding 95% confidence interval were to be reported. The stratified log-rank test was used to test the null hypothesis that the distributions for the time to recovery of normal sensation of the lip were equal between the 2 treatment groups vs. the alternative hypothesis that the distributions were different. The location of the dental procedure (mandibular and maxillary) stratification factor was used for computing the stratified log-rank test statistic.
>
> The hazard ratios for the treatment groups were estimated from a stratified proportional hazards model, with treatment group and location of procedure (mandible or maxilla) included as fixed effects. The consistency of treatment across strata was evaluated by testing for the presence of a

treatment group-by-strata interaction. If no interaction was detected ($p>0.05$), then the single coefficient for treatment group that was estimated from the stratified model was to be reported. Otherwise, if an interaction was detected, the stratum specific coefficients were to be reported. The adequacy of the model was evaluated, including graphical and analytical assessments of the proportional hazards assumption.

As a secondary analysis, a Weibull accelerated failure time (AFT) model was used to further describe the effect of study treatment as measured by this endpoint. Under the AFT model, the time to recovery of normal sensation of the lower lip for a given subject who received NV-101 was taken to be a multiple of the time required for a subject who received the sham injection. In this case, NV-101 was anticipated to "speed up" the passage of time. The degree by which NV-101 speeded the passage of time was to be based on the acceleration factor estimated from this model. The above methods of analysis were also used on the time to recovery of normal sensation of the tongue.

Patient Disposition, Demographic and Baseline Characteristics

The primary efficacy endpoint was time to observed recovery of normal sensation in the lip as measured by pediatric subjects 6 to 11 years of age that were trainable in the standardized lip palpation procedure and able to perform this assessment. Therefore, the number of subjects in the mITT analysis set was less than the number of randomized subjects (72 NV-101 subjects and 43 sham subjects were included in the mITT analysis set for lip sensation out of 152 randomized subjects). For the mITT lip sensation analysis set, a total of 37 subjects (24 in the NV-101 group and 13 in the sham group) were either 4 to 5 years old or 6 to 11 years old and were not trainable in the standardized palpation procedure. These 37 subjects were excluded from this analysis set and were to be analyzed for safety only (Table 10):

Table 10. Patients' Accountability N (%), (Safety)

	NV-101		***Sham***	
	Mandible	***Maxilla***	***Mandible***	***Maxilla***
Randomized patients	49	47	26	30
Completed treatment period	49	47	26	30
Discontinued	0	0	0	0
Analysis Population				
MITT – Lip	38 (77.6)	34 (72.3)	19 (73.1)	24 (80)
MITT – Tongue	32 (65.3)	0	16 (61.5)	0
Safety	49	47	26	30
Number of ITT patients with a protocol deviation	13 (26.5)	15 (31.9)	7 (26.9)	7 (23.3)
Inclusion/Exclusion criteria	1 (2.0)	2 (4.3)	0	0
Study drug	2 (4.1)	1 (2.1)	0	0
Randomization	0	1 (2.1)	1 (3.8)	0
Study procedure	11 (22.4)	11 (23.4)	6 (23.1)	6 (20.0)
Blinding	0	3 (6.4)	0	1 (3.3)

* Results from reviewer's analysis; code: demo.sas.

Demographic characteristics are summarized in Table 11 for the safety population. The treatment groups were well balanced for race, ethnicity, age grade, height, and weight. Nearly equal numbers of males and females were enrolled. About half of the subjects in each treatment group were white. The mean age was approximately 8 years. For those children whose grade in school was known, enrollment ranged from kindergarten through sixth grade in both groups.

Table 11. ITT Subjects' Demographics and Baseline Characteristics by Treatment

	NV-101 (N=96)			Sham (N=56)		
	Mandible N=49	Maxilla N=47	Total N=96	Mandible N=26	Maxilla N=30	Total N=56
Sex, n (%)						
Male	24 (49.0)	19 (40.4)	43 (44.8)	14 (53.8)	18 (60.0)	32 (57.1)
Female	25 (51.0)	28 (59.6)	53 (5.2)	12 (46.2)	12 (40.0)	24 (42.9)
Race, n (%)						
White	25 (51.0)	25 (53.2)	50 (52.1)	16 (61.5)	14 (46.7)	30 (53.6)
Black	14 (28.6)	15 (31.9)	29 (30.2)	5 (19.2)	10 (33.3)	15 (26.8)
Asian	3 (6.1)	2 (4.3)	5 (5.2)	0	3 (10/0)	3 (5.4)
Native Hawaiian	0	0	0	1 (3.8)	0	1 (1.8)
Other	7 (14.3)	5 (10.6)	12 (12.5)	4 (15.4)	3 (10.0)	7 (12.5)
Age Group, n (%)						
4-5	7	9	16	4	2	6
6-11	42	38	80	22	28	50
Mean (SD)	7.9 (2.1)	7.7 (2.0)	7.8 (2.0)	7.7 (2.2)	7.7 (2.2)	7.7 (2.2)
Median	8	8	8	8	8	8
Range	4 – 11	4 – 11	4 – 11	4 – 11	4 – 11	4 – 11
Primary injection type						
Inferior alveolar nerve block	.	.	28 (29.2)	.	.	16 (28.6)
Mental-incisive block	.	.	4 (4.2)	.	.	2 (3.6)
Gow-Gates nerve block	.	.	12 (12.5)	.	.	8 (14.3)

* Results from reviewer's analysis; code: demo.sas.

Results and Conclusions

The primary efficacy endpoint, time to the observed recovery of normal sensation in the lip (combined data for upper and lower lip), is summarized in Table 12.

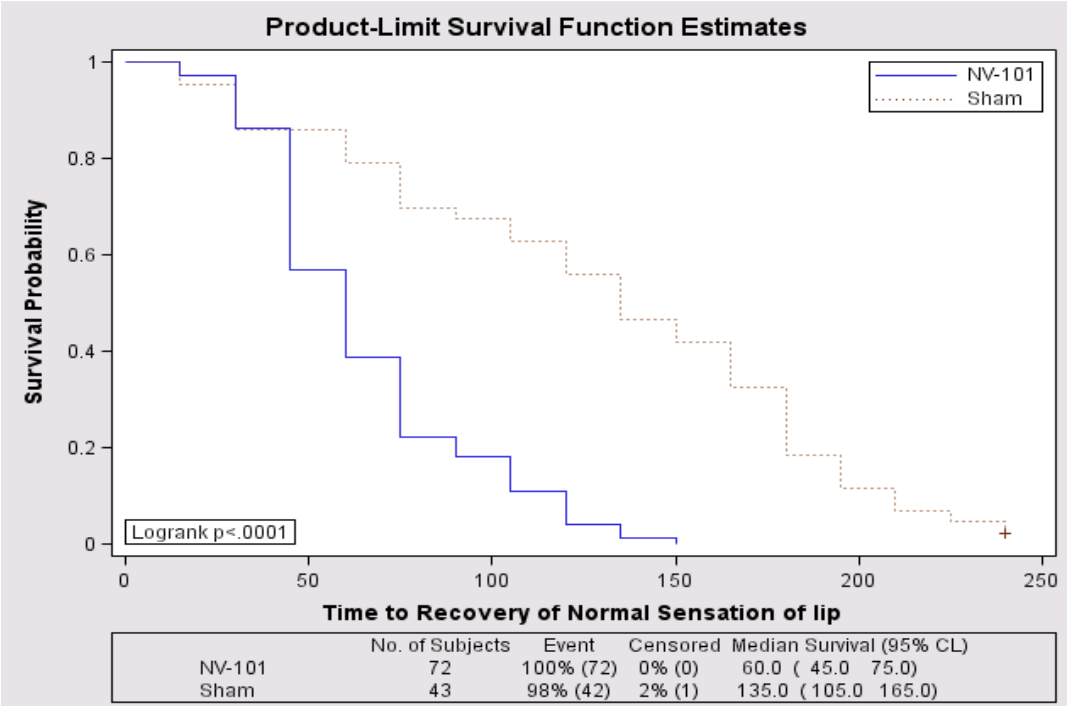
The estimated median time to normal sensation of the lip was 60 minutes for the subjects treated with NV-101 and 135 minutes for the subjects treated with sham. One subject (300-09-011), who did not reach normal lip sensation by the end of the 4-hour observation period, was censored in the analysis. There was a statistically significant difference in time to recovery of normal sensation of the lip between NV-101 and sham as analyzed by the log-rank test stratified by the location of dental procedure (mandibular or maxillary). The reduction in median time to normal sensation of the lip is illustrated below in Figure 7 and Figure 8.

Table 12. Log-Rank Analysis of Time to Recovery of Normal Sensation in Lower (or Upper) Lip

	Lip		Tongue	
	NV-101 (n=72)	Sham (n=43)	NV-101 (n=32)	Sham (n=16)
Median Time to recovery of normal sensation[1] (min.)	60	135	45	112.5
95% CI	(45, 75)	(105, 165)	(30, 45)	(45, 150)
NV-101 vs. placebo				
Hazard ratio[2]		5.90		5.27
95% CI		(3.35, 10.4)		(2.04, 13.7)
p-value		<0.001		<0.001

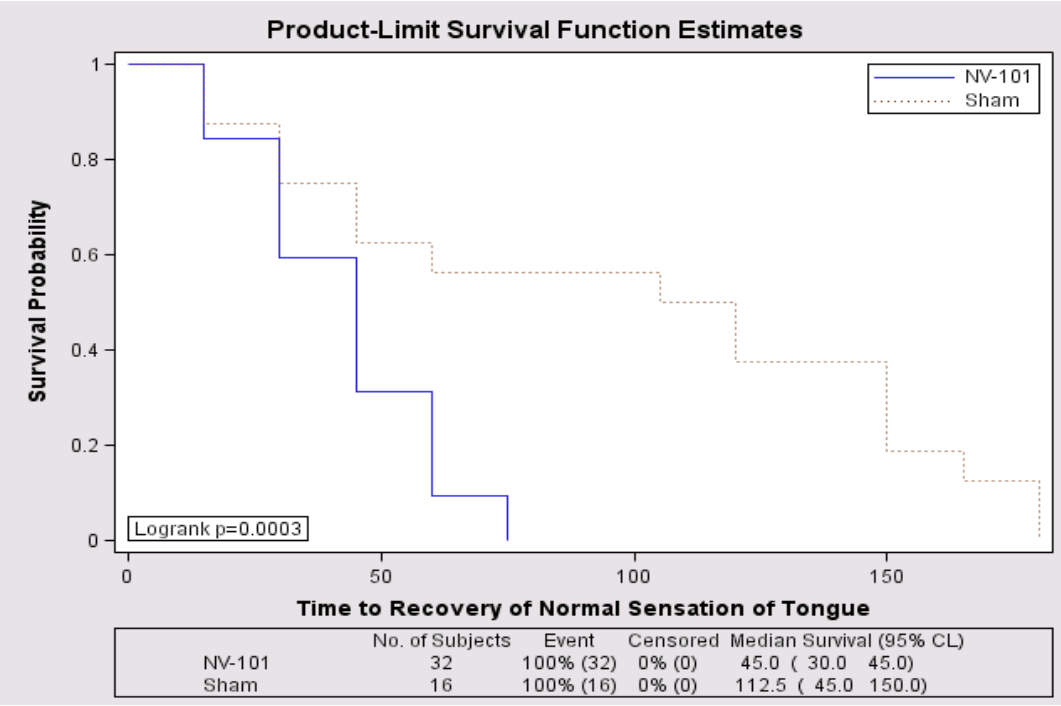
Note: 1. Kaplan-Meier estimated; 2. Log-Rank test stratified for the location of dental procedure.
* Results from reviewer's analysis; code: km_anal.sas.

Figure 7. Kaplan-Meier Plot of Time to Recovery of Normal Sensation in Lip (MITT)



* Results from reviewer's analysis; code: km_anal.sas.

Figure 8. Kaplan-Meier Plot of Time to Recovery of Normal Sensation in Tongue (MITT)



* Results from reviewer's analysis; code: km_anal.sas.

The Cox proportional hazards model was employed to estimate the hazard ratio for the treatment group. The sponsor's model included effects of treatment group and location of dental procedure (mandible, maxilla). The consistency of treatment across strata was evaluated by testing for the presence of treatment group-by-stratum interaction. Because these interaction effects were not statistically significant, the interaction terms were dropped from further Cox proportional hazards modeling. I additionally analyzed a model with effects for group, location of dental procedure, and elapsed time.

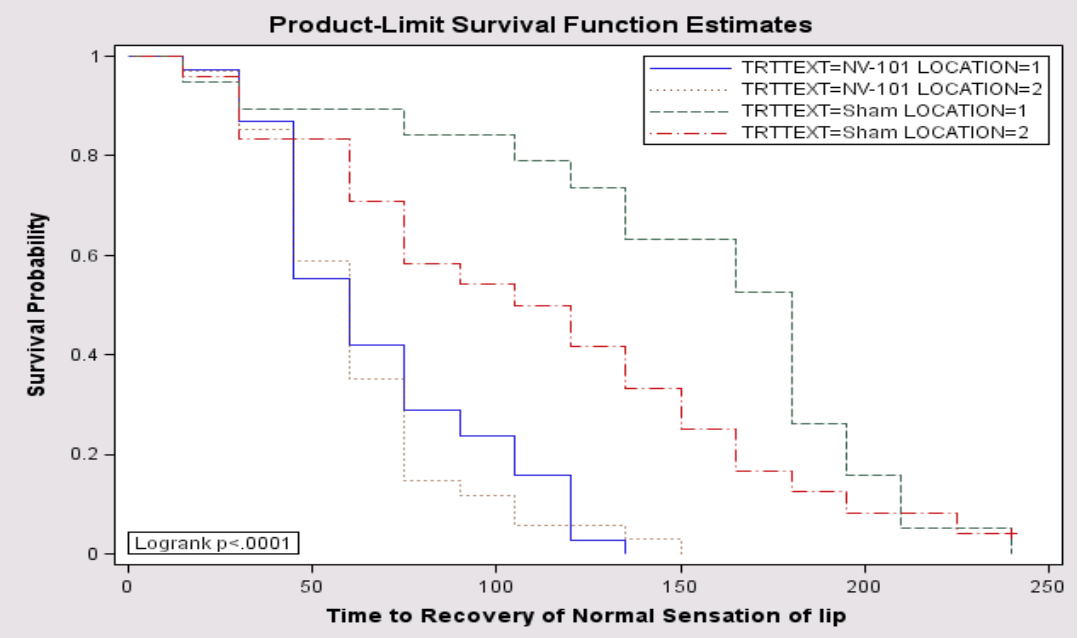
The results of my analysis are shown in Table 13. The hazard ratio for the treatment group in the model was 4.16 indicating that after adjusting for location of dental procedure and elapsed time, subjects in the NV-101 treatment group were 4 times as likely as subjects in the sham group to achieve normal lip sensation during the 4-hour observation period. The lack of statistical significance for the location of dental procedure indicates that there was no difference between the procedures on time to recovery of normal lip sensation. This effect is further illustrated by a Kaplan-Meier plot of the data, as shown in Figure 9.

Table 13. Cox Proportional Hazards Model for Time to Recovery of Normal Sensation

Variable	***Lip***		***Tongue***	
	Hazard Ratio	**p-value 95%CI**	**Hazard Ratio**	**p-value 95%CI**
Treatment (NV-101 vs. sham)	4.16	<0.0001 (2.53, 6.83)	3.51	0.003 (1.51, 8.34)
Location (Mandible vs. Maxilla)	0.78	0.190 (0.54, 1.13)	.	.
Elapsed time (min) (≤ 32 vs. > 32)	0.55	0.109 (0.27, 1.14)	0.90	0.735 (0.51, 1.62)

* Results from reviewer's analysis; code: km_anal.sas.

Figure 9. Kaplan-Meier Plot of Time to Recovery of Normal Sensation in Lower Lip (ITT)



* Results from reviewer's analysis; code: km_anal.sas.

3.2 Evaluation of Safety

The evaluation of the safety data was conducted by Dr. Arthur Simone. The reader is referred to Dr. Simone's review for information regarding the adverse event profile.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Others

Consistent with results for the overall cohort, differential distributions in recovery times between the randomized treatment groups were also apparent in subsets of subjects categorized based on number of cartridges received, type of anesthetic, age group, type of dental procedure, type of injection, and sex." Kaplan-Meier analyses of these subsets demonstrated that the ability of NV-101 to shorten the time required for recovery of normal sensation in the lower lip was observed for subjects treated with either 1 or 2 cartridges/sham injections, for subjects in all 3 age groups, for subjects treated with either inferior alveolar block or mental-incisive block, for subjects undergoing cavity preparation/restoration/filling or periodontal maintenance, and for both males and females. The effect of NV-101 appeared to be consistent across all of the subgroups analyzed, with the reduction factors ranging from 45.3% to 70.8%.

Table 14. Subgroup Analysis of Time to Recovery of Normal Sensation in the Lip (Two Studies)

Variable	*NV-101*		*Sham*		% Reduction
	N	Median Time (minutes)	N	Median Time (minutes)	Log-rank test p-value
Overall	242	62.5	242	140	55.4% (p<0.0001)
Sex					
Male	122	65	109	145	55.2% (p<.0001)
Female	120	60	133	140	57.1% (p<.0001)
Age Group					
12 to 17 years	26	87.5	29	160	45.3% (p=0.01)
18 to 64 years	187	60	187	140	57.1% (p<.0001)
≥ 65 years	29	55	26	112.5	51.1% (p<.0001)
Race					
White	191	65	186	140	53.6% (p<.0001)
Non-White	51	60	56	152.5	60.7% (p<.0001)
Number of Cartridges					
1	204	57.5	207	140	58.9% (p<.0001)
2	38	85	35	155	45.2% (p<.0001)
Anesthetic					
Lidocaine	161	60	161	140	57.1% (p<.0001)
Other	81	75	81	155	51.6% (p<.0001)
Dental Procedure					
Cavity	166	67.5	167	145	53.4% (p<.0001)
Crown	3	30	7	130	76.9% (p<.001)
Periodontal maintenance	73	55	68	142.5	61.4% (p<.0001)
Type of Injection					
Inferior Alveolar Nerve Block	201	70	199	145	51.7% (p<.0001)
Other	41	35	43	120	70.8% (p<.0001)

* Results from reviewer's analysis; code: sub_group.sas.

4.2 Other Special/Subgroup Populations

There are no other special/subgroup analyses.

5. SUMMARY AND CONCLUSIONS

Novalar Pharmaceuticals has proposed OraVerse® (phentolamine mesylate) injection for "the reversal of soft tissue anesthesia and the associated functional deficits resulting from an intraoral submucosal injection of a local anesthetic containing a vasoconstrictor." The sponsor conducted seven studies including three phase 3 studies, referenced in the label, to evaluate the efficacy of NV-101.

Based on my collective evaluation of the NDA submission, I conclude that there is evidence of the effect of phentolamine mesylate in reversal of STA in subjects undergoing dental procedures involving the mandible and maxilla with an anesthetic/vasoconstrictor combination. Studies demonstrated the faster recovery of normal lip sensation, as well as normal abilities to smile, speak, drink refrain from drooling.

5.2.1 Labeling

The sponsor's draft label references two adult studies and one pediatric study. I recommend the following changes:

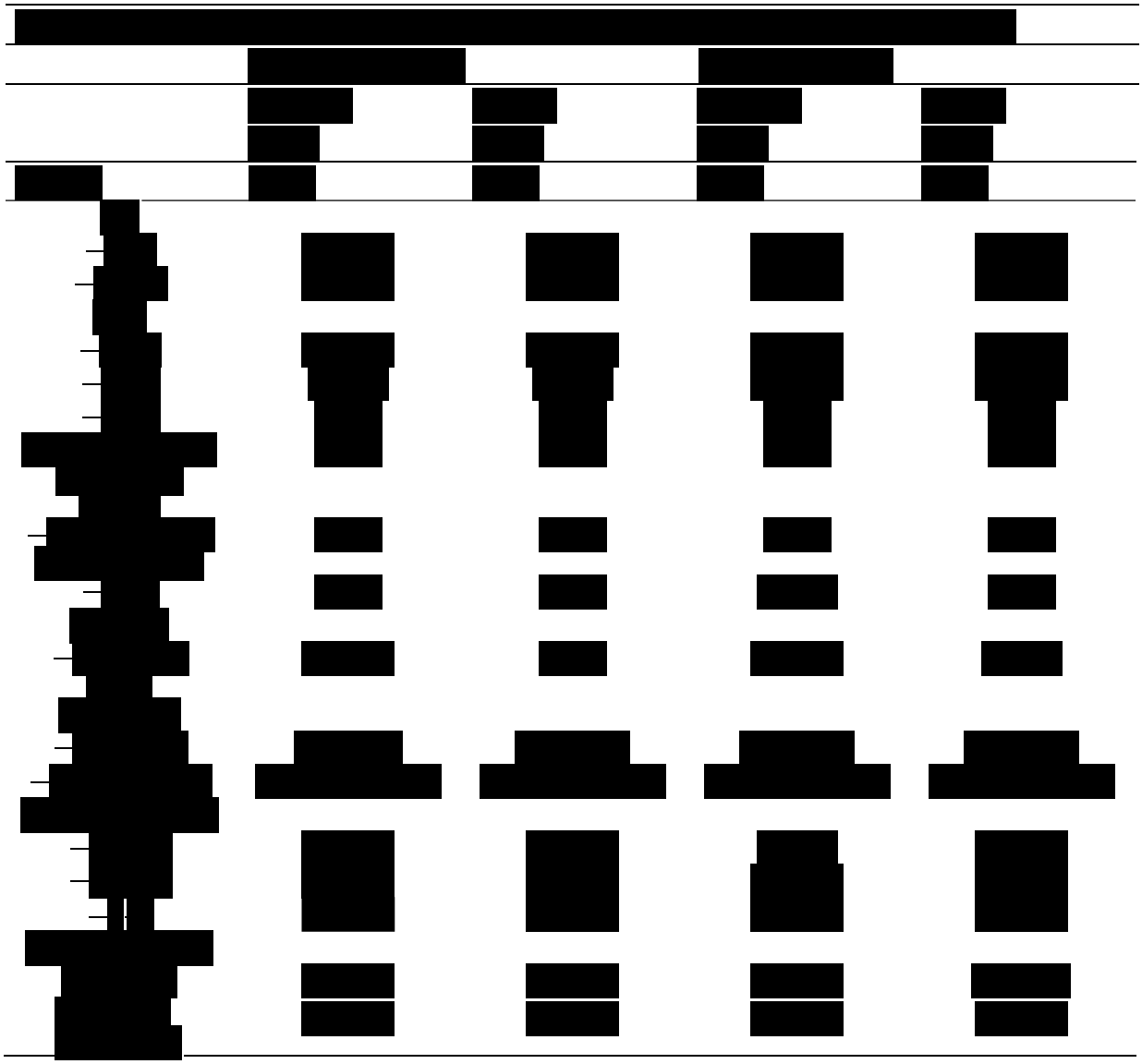
14. CLINICAL STUDIES

The safety and efficacy of OraVerse in the treatment of soft tissue anesthesia (STA) following a dental procedure that required local anesthesia containing a vasoconstrictor were studied in the following [redacted] clinical studies.

[redacted] Phase 3 blinded, randomized, multi-center controlled studies conducted in dental patients who had mandibular (Study 1, [redacted] or maxillary (Study 2, [redacted]) restorative or periodontal maintenance procedures and had received a [redacted] anesthetic that contained a vasoconstrictor. The primary endpoint was time to normal lip sensation as measured by patient reported lip [redacted]. The secondary endpoints [redacted] patient's perception of altered function, sensation and appearance, and actual functional deficits in smiling, speaking, drinking and drooling as [redacted] by both the patient and an observer blinded to the treatment. [redacted] In the mandibular study, the time to recovery of tongue sensation was also a secondary endpoint. [redacted] Patients were stratified by type and amount of anesthetic administered [redacted]

[redacted]
OraVerse was administered at a cartridge ratio of 1:1 to local anesthetic.

[redacted]



[redacted] The median time to recovery of normal sensation in the lower lip was reduced by 85 minutes [redacted]. The median time to recovery of normal sensation in the upper lip was reduced by [redacted] minutes (62[redacted]%).

Patients randomized to OraVerse or control recovered in 50 minutes and 132.5 minutes, [redacted] Kaplan-Meier [redacted] for time to normal lip sensation. [redacted]

Figure 1: Kaplan-Meier Plot of Time to Recovery of Normal Sensation in the Lower Lip (ITT Analysis Data Set)

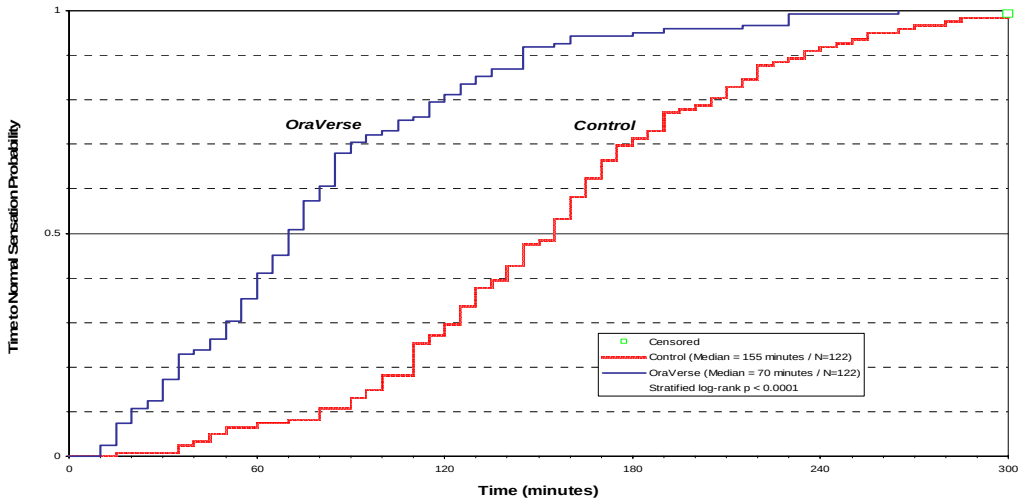
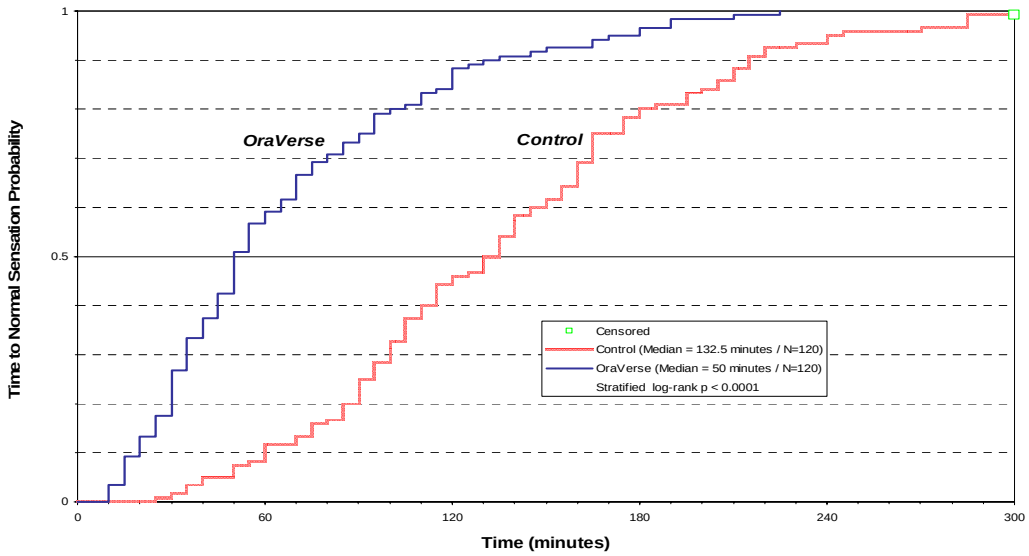


Figure 2: Kaplan-Meier Plot of Time to Recovery of Normal Sensation in the Upper Lip (ITT Analysis Data Set)



[redacted] In Study 1 (mandibular), OraVerse accelerated: a) the recovery of the perception of normal appearance and function [redacted] by 60 minutes (40%), b) the recovery of normal function [redacted] by ~~60~~ [redacted] minutes (50%), and c) the recovery of normal sensation in the tongue by 65 minutes (52%). In Study 2 (maxillary), the [redacted] was reduced by 60 minutes (50%) and [redacted] by ~~45~~ [redacted] minutes [redacted].

[redacted]

Study 3, a Pediatric Phase 2, blinded, randomized, multi-center controlled study was conducted in dental patients who had received 2% lidocaine with 1:100,000 epinephrine. Dental patients (n=152, ages 4-11 years) received 1/2 cartridge of local anesthetic if weighing ≥15 kg and <30 kg and one-half or one full cartridge if weighing ≥30 kg [redacted] at a cartridge ratio of 1:1 to local anesthetic.

The [redacted] time to normal lip sensation, [redacted] in [redacted] patients 6 to 11 years of age who were trainable in the lip [redacted] procedures [redacted] was reduced by 75 minutes [redacted]

-EOF-

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

/s/

Feng Zhou
12/7/2007 02:11:37 PM
BIOMETRICS

Dionne Price
12/7/2007 02:18:26 PM
BIOMETRICS
Concur