CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

125274

STATISTICAL REVIEW(S)

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

BIOLOGICS LICENSING APPLICATION

CLINICAL STUDIES

BLA/Serial Number:	125286
Drug Name:	Reloxin (botulinum type A toxin)
Indication(s):	Glabellar lines
Applicant:	Ipsen / Medicis
Dates:	Submitted: 3/12/08 PDUFA: 1/12/09
Review Priority:	Standard review
Biometrics Division:	Division of Biometrics III
Statistics Reviewer:	Kathleen Fritsch, PhD
Concurring Reviewer:	Mohamed Alosh, PhD
Medical Division:	Division of Dermatology and Dental Products
Clinical Team:	Denise Cook, MD
Project Manager:	Tamika White

Keywords: randomization issues, dose ranging, botulinum toxin

Table of Contents

1 EXECUTIVE SUMMARY 4

1.1 Conclusions and Recommendations 4

1.2 Brief Overview of Clinical Studies 5

1.3 Statistical Issues and Findings 5

2 INTRODUCTION 6

2.1 Overview 6

2.2 Data Sources 9

3 STATISTICAL EVALUATION 9

3.1 Evaluation of Efficacy 9
3.1.1 Study Y-97-52120-719 10
3.1.1.1 Study Design 10
3.1.1.2 Subject Disposition 11
3.1.1.3 Baseline Characteristics 11
3.1.1.4 Day 30 Efficacy Results 12
3.1.1.5 Efficacy by Center 13
3.1.1.6 Duration of Effect 14
3.1.2 Study Y-97-52120-085 15
3.1.2.1 Study Design 15
3.1.2.2 Subject Disposition 17
3.1.2.3 Baseline Characteristics 18
3.1.2.4 Day 30 of Cycle C Efficacy Results 19
3.1.2.5 Efficacy by Center 20
3.1.2.6 Efficacy over Time 20
3.1.2.7 Randomization Issues 21
3.1.3 Study A-2006-01 23
3.1.3.1 Study Design 23
3.1.3.2 Subject Disposition 24
3.1.3.3 Baseline Characteristics 25
3.1.3.4 Day 30 Efficacy Results 26
3.1.3.5 Efficacy Results by Center 27
3.1.3.6 Efficacy over Time 28
3.1.3.7 Variable Dose 31
3.1.4 CAMR Product Studies 32
3.1.4.1 Study Y-97-52120-096 32
3.1.4.2 Study Y-97-52120-718 33

3.2 Evaluation of Safety 38

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS 41

4.1 Gender, Race, and Age 41

4.2 Other Special/Subgroup Populations 44

5 SUMMARY AND CONCLUSIONS 45

5.1 Statistical Issues and Collective Evidence 45
5.1.1 Fixed versus Variable Dosing 45
5.1.2 Efficacy of 50 U Dose 45
5.1.3 Statistical Issues 46

5.2 Conclusions and Recommendations 47

SIGNATURES/DISTRIBUTION LIST 48

1 Executive Summary

1.1 Conclusions and Recommendations

The applicant has demonstrated the efficacy of Reloxin (50 U [units]) in the treatment of glabellar lines relative to placebo in two studies (719 and 085). This review did not identify any significant statistical concerns regarding the design or conduct of Study 719 that would impact the efficacy conclusions. Study 085 had a problem with its initial randomization, which required the study to be modified with an additional treatment cycle. Subjects who had received Reloxin in all previous treatment cycles were eligible to be randomized into the final added cycle. Although the study design was modified, all modifications were completed before subjects were randomized into the final treatment cycle, and thus the results are interpretable.

The applicant has also demonstrated that a variable dose of Reloxin (50-80 U) with dosing based on gender and the investigator's visual assessment of procerus and corrugator muscle mass is efficacious relative to placebo in one study (2006-01). However, the applicant has not provided sufficient evidence that the higher doses provide greater efficacy (no comparative dose-ranging information has been provided), and a higher rate of ptosis was observed in female subjects who received the highest dose for females (70 U). No long-term safety data on repeat dosing with doses higher than 50 U has been submitted. Therefore the applicant has not submitted sufficient dose-ranging and safety information to justify that there is any efficacy/risk benefit to the higher dose.

In addition to the three efficacy studies conducted on the to-be-marketed version of Reloxin, the applicant conducted a study on an earlier version of Reloxin made from an alternative bulk active substance (referred to as CAMR product). Study 718 demonstrated that CAMR Reloxin was superior to placebo.

Efficacy endpoints in Studies 719 and 085 were based on the investigator and subject assessments at maximum frown on Day 30. The investigator and subject assessments were 4-point scales with categories: none, mild, moderate, and severe. All subjects were classified as moderate or severe on both the investigator and subject assessments at baseline. The protocol-specified co-primary endpoints were '1+ response' on the investigator and subject scales, defined as achieving scores of 'none' or 'mild' on Day 30. In Study 085 the endpoints were assessed on Day 30 in Cycle C. In addition, the studies also demonstrated efficacy on response definitions of interest to the Agency: 1+ composite response (none or mild on *both* the investigator and subject assessments), and 2+ composite response (none or mild with at least 2 grades reduction from baseline on both the investigator and subject assessments). The 2+ composite response rates are presented in Table 1. The totality of evidence from the clinical studies supports the claim that Reloxin (50 U) has demonstrated efficacy in the treatment of glabellar lines.

Table 1 – 2+ Composite Response Rates in Phase 3 Studies

Study	Dose/Product	Reloxin	Placebo	P-value
719	50 U	58/105 (55%)	0/53 (0%)	<0.001
085	50 U	37/71 (52%)	0/71 (0%)	<0.001
2006-01	50 – 80 U	319/544 (59%)	1/272 (<1%)	<0.001
718	50 U (CAMR)	120/200 (60%)	0/100 (0%)	<0.001

1.2 Brief Overview of Clinical Studies

The applicant has conducted 8 clinical studies with Reloxin in the treatment of glabellar lines including a dose-ranging study, four Phase 3 studies, two open-label safety studies, and a comparability study. Refer to Table 2 in Section 2.1 for a complete description of the clinical studies program. This review focuses on the four Phase 3 studies. All studies were conducted in the United States.

1.3 Statistical Issues and Findings

Study 719 demonstrated that 50 U of Reloxin is superior to placebo in the treatment of glabellar lines. This review did not identify any significant statistical concerns regarding the design or conduct of Study 719 that would impact the efficacy conclusions.

The original design of Study 085 was to treat subjects with one or two open-label treatment cycles of Reloxin (Cycles A1 and A2) and then randomize subjects to either Reloxin (50 U) or placebo in Cycle B and assess efficacy on Day 30 of Cycle B. However, the randomization in Cycle B failed and subjects were allocated treatments sequentially, that is, all vials with Reloxin were dispensed before placebo (within age strata). Thus the applicant unblinded the data from Cycle B, added an additional treatment cycle (Cycle C), identified subjects who had received Reloxin in Cycle B, re-consented eligible and willing subjects and randomized them to either Reloxin or placebo in Cycle C. In Study 085 the failed randomization is a significant concern. However, because the goal of the study was to assess efficacy after a few lead-in cycles of Reloxin, adding an additional cycle may have impacted the selection of subjects who entered the final randomized cycle, but otherwise the study was able to evaluate a group of randomized subjects and assess efficacy. The efficacy results in Study 085 are consistent with the results from other Reloxin studies.

Study 2006-01 was able to demonstrate that a variable dose of Reloxin is superior to placebo. Most subjects in Study 2006-01 received either 60 or 70 U of Reloxin. The applicant has requested labeling for the variable dosing strategy. The only safety information on doses greater than 50 U comes from Study 2006-01. All of the submitted data from the long-term safety studies used the 50 U dose of Reloxin. Although Study 2006-01 was the largest efficacy study, it was a single-dose study, and no repeat-dose information on doses higher than 50 U has been submitted. Although it has been consistently observed across studies that males tend to have lower response rates than females, the applicant has not provided convincing evidence that higher doses of Reloxin lead to higher response rates. No within-study comparative data have been submitted, and

other studies that treated males with 50 U of Reloxin had similar response rates in males as Study 2006-01. Higher levels of ptosis were observed with 70 U than 60 U of Reloxin in Study 2006-01 (3.8% vs. 1.4%). Thus although the applicant has demonstrated that the variable dosing scheme is an effective dose, the applicant has not presented convincing evidence that there is an efficacy benefit to the higher doses over the fixed dose and that any efficacy benefit justifies possible increases in safety events such as ptosis.

Although most subjects received a higher dose of Reloxin, the response rates in Study 2006-01 were fairly comparable to those observed in other studies. Because the dose response curve does not appear to be very steep within the range of 50 to 80 U, Study 2006-01 provides additional support that Reloxin is efficacious, even though relatively few subjects in the study received 50 U, and it cannot be directly inferred that efficacy at a higher dose implies efficacy at a lower dose.

Study 718 demonstrated the efficacy of CAMR Reloxin. The small comparability Study 096 provides some additional support that to-be-marketed product is not substantially less effective than CAMR product, although this study was not designed to establish equivalence between the products. Thus, Study 718 may also provide supportive evidence of efficacy for Reloxin.

2 Introduction

2.1 Overview

Reloxin (botulinum type A toxin) is a neurotoxin for intramuscular injection for the treatment of glabellar lines. The same product, under the name Dysport, is currently under review in the Division of Neurology Products for the treatment of cervical dystonia. The applicant has conducted a number of Phase 2, Phase 3, and long-term safety studies. The earliest studies used product formulated from bulk active substances (BAS) manufactured at the Centre for Applied Microbiology and Research (CAMR). The later studies used product formulated from BAS manufactured at Ipsen Biopharm Limited (IBL). The IBL product is the to-be-marketed product. The CAMR product and IBL product have been noted to have some analytical differences.

Most of the clinical development program used a fixed dosing regimen of 50 U (units) per treatment session. However, the final study conducted by the applicant (2006-01) used a variable dosing scheme of 50, 60, or 70 U for women and 60, 70, or 80 U for men based on clinical assessment of procerus and corrugator muscle mass (small, medium, or large). All of the long-term repeat-dose studies used the fixed 50 U dosing regimen. The eight studies in the applicant's clinical development program are outlined in Table 2. All studies were conducted in the United States. This review will focus on Studies 719, 085, 2006-01, and 718. This review will briefly discuss Study 096.

Table 2 – Clinical Study Program for Reloxin

Study	*Subjects*	*Dates*	*Comments*
717	Placebo – 94 Reloxin (C) 20 U – 90 Reloxin (C) 50 U – 95 Reloxin (C) 75 U – 94	12/2002 – 7/2003	Phase 2 dose ranging study. Used CAMR product. Single dose.
718	Placebo – 100 Reloxin (C) 50 U – 200	4/2004 – 12/2004	Phase 3 study. Used CAMR product. Single dose.
720	Reloxin 50 U: Cycle 1 (C) – 1200 Cycle 2 (C/I) – 1145 Cycle 3 (I) – 1031 Cycle 4 (I) – 661 Cycle 5 (I) – 177	10/2004 – 3/2006	Long-term safety study. Used CAMR and IBL product. Multiple dose (up to 5 doses).
096	Reloxin (C) 50 U – 50 Reloxin (I) 50 U – 50	2/2005 – 4/2005	Comparability of CAMR and IBL product. Single dose.
085	Cycle A1 Reloxin (I) 50 U – 311 Cycle A2 Reloxin (I) 50 U – 190 Cycle B Reloxin (I) 50 U – 171 Cycle C Placebo – 71 Reloxin (I) 50 U – 71	6/2005 – 4/2007	Phase 3 study. Used IBL product. Multiple dose (up to 4 doses).
719	Placebo – 53 Reloxin (I) 50 U– 105	11/2005 – 7/2006	Phase 3 study. Used IBL product. Single dose.
732	Reloxin 50 U: Cycle 1 (I) – 768 Cycle 2 (I) – 607 Cycle 3 (I) – 470 Cycle 4 (I) – 284 Cycle 5 (I) – 123 Cycle 6 (I) – 7	11/2005 – Ongoing (Data cutoff 3/2007)	Long-term safety study. Used IBL product. Subjects rolled over from Studies 085, 719, 720, or 2006-01. Multiple dose. Up to 8 doses can be given in this study; additional dose(s) would have been given in previous study.
2006-01	*Females* Placebo – 238 Reloxin (I) 50 U – 22 Reloxin (I) 60 U – 277 Reloxin (I) 70 U – 181 Reloxin (I) 80 U – 1 *Male* Placebo – 34 Reloxin (I) 60 U – 5 Reloxin (I) 70 U – 25 Reloxin (I) 80 U – 33	12/2006 – 7/2007	Phase 3 study. Used IBL product. Single dose. Variable dosing based on muscle mass.

Note: (C) = CAMR product, (I) = IBL product

The applicant's development program occurred in several stages. In the first stage the studies used the CAMR product. The sponsor conducted a Phase 2 dose-ranging study (717), a Phase 3 efficacy study (718), and initiated a long-term safety study (720) using the CAMR product. After this, the sponsor switched to the IBL product. The sponsor conducted a bridging study of the IBL and CAMR product (096) to support using the same dose level (50 U) selected from the dose-ranging study.

Next, the sponsor initiated two Phase 3 studies (085 and 719) and a long-term safety study (732) using the IBL product. These three studies, which all used the 50 U dose, were intended to provide the majority of efficacy and safety information to support a BLA. However, during the conduct of Study 085, the applicant realized that there was a problem with the randomization in this study. A subsequent audit revealed that subjects in the 'randomized' cycle of Study 085 were receiving treatment assignments sequentially, rather than randomly. That is, all doses of Reloxin were assigned within a stratum before doses of placebo. The sponsor revised the protocol of Study 085 adding an additional randomized treatment cycle to subjects remaining in the study. However, FDA recommended that the applicant conduct an additional study, as the randomization problems from Study 085 could make the study inadequate as a pivotal study.

Thus, the sponsor proposed Study 2006-01. This study differed from all of the previous studies in that subjects received a variable dosing scheme of 50 to 80 U based on gender and perceived muscle mass rather than the fixed dosing scheme of 50 U. The applicant has not conducted any long-term repeat-dose studies of the variable dosing scheme. Subjects from Study 2006-01 are eligible to rollover into Study 732, and the sponsor has amended the protocol for Study 732 to allow subjects rolling over from Study 2006-01 to continue receiving the dose the were assigned in Study 2006-01. However, as of the cutoff date for the data submission, no subjects from Study 2006-01 had yet provided any repeat dose data.

The applicant has requested labeling for the variable dosing scheme (50 to 80 U). The applicant has hypothesized that the larger muscle mass size would require higher doses of Reloxin. However, the applicant has not conducted any dose ranging by muscle mass size to see if this is in fact the case. It is not clear from the submission how accurately investigators could assess the muscle mass and whether muscle mass is truly an important factor for determining the appropriate dose. Study 2006-01 enrolled a limited number of subjects at the highest dose (only 34 subjects received 80 U). The applicant has not provided any long-term safety data on the variable dosing scheme. Long-term data on a lower dose would not be sufficient to ensure safety for a higher dose.

Due to the lack of data to support the long-term safety and the lack of information that efficacy varies due to gender and muscle mass, a single Phase 3 study demonstrating that a variable dose of Reloxin based on gender and muscle mass is superior to placebo is not sufficient to support labeling for the variable dose. Therefore this review will focus on assessing whether the completed studies provide adequate information to support the efficacy and safety of 50 U of Reloxin in the treatment of glabellar lines.

This review will evaluate four studies for efficacy: 719, 085, 2006-01, and 718. Brief results from Study 096 will also be presented. Three of these studies have serious issues that will need to be addressed regarding the studies' ability to demonstrate the efficacy of Reloxin 50 U for the treatment of glabellar lines. These concerns include

- Study 085 had a randomization failure that required the study to be modified with an additional randomized treatment cycle during the study
- Study 2006-01 uses a variable dosing strategy rather than the fixed 50 U dose
- Study 718 used CAMR rather than IBL product. The applicability of the data from this study will depend on whether the CAMR and IBL product are clinically comparable.

2.2 Data Sources

This reviewer evaluated the sponsor's clinical study reports and clinical summaries, as well as the proposed labeling. This submission was submitted in eCTD format and was entirely electronic. The datasets used in this review are archived at \\Cbsap58\M\eCTD_Submissions\STN125286.

3 Statistical Evaluation

3.1 Evaluation of Efficacy

The applicant conducted four placebo-controlled efficacy studies to support the efficacy of Reloxin in the treatment of glabellar lines. Due to differences in design, each study will be evaluated separately. However, in each study, efficacy was assessed using the same investigator and subject assessment scales. The investigator's scale used the following categories: 0 = none, 1 = mild, 2 = moderate, 3 = severe, and investigators were provided with a set of four reference photographs, one for each grade. The subject's scale also had four categories, but did not involve reference photographs. The subject's scale used the descriptors 0 = no wrinkles (smooth skin), 1 = mild wrinkles (fairly smooth skin), 2 = moderate wrinkles (glabellar lines), and 3 = severe wrinkles (deep glabellar lines).

To unify response definitions across studies, the applicant defined two response definitions: '1+ response' and '2+ response'. A 1+ response is defined as achieving a score of 0 or 1 on the given scale. A 2+ response is defined as achieving 0 or 1 with at least 2 grades reduction from baseline. In each definition of response, a score of 0 or 1 is required. Note in particular that the definition of 1+ response is *not* equivalent to 'at least 1 grade reduction from baseline', as a subject with a baseline score of 3 must reach a score of 0 or 1 to become a responder. Although these terms may be somewhat confusing, this review will use the same terminology as the application.

Except where noted differently, the applicant handled missing data for the response rate analyses in two ways. In the first method, values for missing assessment scores were imputed using the average of all non-missing scores (both treatments combined, rounding to the nearest integer) for the two treatment groups. Thus for 1+ response, if the average of all observed scores was less than 1.5, all subjects with missing data would be imputed as responders, and if the average was greater than 1.5 all subjects with missing data

would be imputed as non-responders. In the second method, missing values were imputed as non-response. In most examples from the clinical studies discussed below, the net effect of the two imputation strategies is that under the first method, subjects with missing values in both arms are treated as responders, while under the second method, subjects with missing values in both arms are treated as non-responders. Where not otherwise specifically noted, this review uses the strategy of treating subjects with missing data as non-responders.

3.1.1 Study Y-97-52120-719

3.1.1.1 Study Design

Study 719 was a randomized, placebo-controlled, double-blind, single dose study to evaluate the efficacy and safety of Reloxin in the treatment of glabellar lines. Subjects were treated with a single dose of either Reloxin (50 U) or placebo and followed for 180 days. The study enrolled 158 subjects at 3 U.S. centers, 105 treated with Reloxin and 53 treated with placebo. Randomization was stratified by age (≤ 50, >50). Subjects were to be naïve to botulinum toxin treatment. Subjects were to be age 18 or older with moderate to severe vertical glabellar lines at maximum frown, as assessed by both the subject and investigator. The investigator's scale used the following categories: 0 = none, 1 = mild, 2 = moderate, 3 = severe, and investigators were provided with a set of four reference photographs, one for each grade. The subject's scale also had four categories, but did not involve reference photographs. The subject's scale used the descriptors 0 = no wrinkles (smooth skin), 1 = mild wrinkles (fairly smooth skin), 2 = moderate wrinkles (glabellar lines), and 3 = severe wrinkles (deep glabellar lines). In addition to the investigator's and subject's assessment, each subject's photographs were evaluated by 3 independent reviewers using the investigator scale.

Subjects completed a diary card on days 1 through 7 to record the onset of treatment. Follow-up clinic visits occurred on Days 14, 30, 60, 90, 120, 150, and 180. The co-primary efficacy endpoints were the investigator's assessment at maximum frown at Day 30 and the subject's assessment at maximum frown at Day 30. For each scale a responder was defined as having a score of 2 or 3 at baseline and a score of 0 or 1 at Day 30 (1+ response). FDA concurred with the choice of endpoints at the End of Phase 2 Meeting held on January 8, 2004. However, FDA requested on April 20, 2004 that responders should have at least a 2 grade reduction from baseline (2+ response). On July 11, 2005, FDA agreed that 2 grades reduction may not be required and that 1+ response may be acceptable. At FDA's request, the applicant also conducted an analysis of composite success from the investigator and subject scores and included the analyses for 2+ response. The protocol included a number of secondary endpoints based on independent reviewer's assessments, investigator assessments, subject assessments, and treatment duration assessments at various timepoints.

The ITT population was defined as all randomized subjects who received study treatment. The per protocol was to exclude subjects with major protocol violations such as use of non-approved medications, violations of inclusion or exclusion criteria, missing or out of window study visits, or incomplete study treatment. No subjects were found to

have had any major protocol violations and the per protocol population was identical to the ITT populations.

Success rates were analyzed with a Cochran-Mantel-Haenszel test stratified on age (≤ 50, > 50). Duration of response was analyzed with a log-rank test stratified on age (≤ 50, > 50). The primary method of handling missing data was to impute the average of all non-missing values (rounded to the nearest integer) for the two treatment groups combined. Thus, if the average score for the Reloxin and placebo groups combined was less than 1.5, then all missing data (in both arms) would be imputed as responders, and if the average score was greater than 1.5, then all missing data would be imputed as non-responders in the 1+ response analyses. As a sensitivity analyses, all missing data would be imputed as non-responders. The applicant did not impute data for the 2+ response analyses.

3.1.1.2 Subject Disposition

Study 719 enrolled 158 subjects. All randomized subjects received their assigned treatment. Two subjects in each arm did not attend the Day 30 visit. Dropout by Day 180 in the studies was similar on both arms with about 10% of the subjects dropping out by Day 180. All dropout was due to either loss to follow-up or subject decision. The subject disposition is presented in Table 3.

Table 3 – Subject Disposition (Study 719)

	Reloxin (50 U)	Placebo
Subjects Randomized	105	53
Attended Day 30 visit	103 (98%)	51 (96%)
Completed Study	97 (92%)	46 (87%)
Discontinuation Reason		
Lost to Follow-Up	3 (3%)	4 (8%)
Subject Decision	5 (5%)	3 (6%)

3.1.1.3 Baseline Characteristics

Baseline demographics were generally balanced across treatment groups. Approximately 85% of the subjects were female. Almost half of the subjects were Caucasian and almost half were Hispanic with a few subjects of other races. The mean age was around 43. The demographic results are presented in Table 4.

Table 4 – Baseline Demographics (Study 719)

	Reloxin (50 U) N=105	Placebo N=53
Age (years)		
Mean	43.1	42.7
Range	(19, 75)	(24, 67)
Gender		
Male	15 (14%)	8 (15%)
Female	90 (86%)	45 (85%)
Race		
Caucasian	52 (50%)	25 (47%)
Hispanic	50 (48%)	25 (47%)
African-American	2 (2%)	0 (0%)
Asian	1 (1%)	2 (4%)
Other	0 (0%)	1 (2%)

In Study 719, baseline severity was balanced across treatment groups. About 60% of subjects were classified as severe on both the investigator and subject assessment scales. See Table 5.

Table 5 – Baseline Assessment at Maximum Frown (Study 719)

	Reloxin (50 U) N=105	Placebo N=53
Investigator		
Moderate	40 (38%)	20 (38%)
Severe	65 (62%)	33 (62%)
Subject		
Moderate	44 (42%)	20 (38%)
Severe	61 (58%)	33 (62%)

3.1.1.4 Day 30 Efficacy Results

Reloxin (50 U) was superior to placebo for both protocol-specified primary efficacy endpoints (1+ response at Day 30 on the investigator and subject assessments). Reloxin was also superior to placebo for the investigator and subject 2+ response assessments, as well as both the 1+ and 2+ composite assessments. Efficacy results are presented in Table 6 and Table 7.

Table 6 – Day 30 Response Rates (1+) at Maximum Frown - ITT (Study 719)

	Reloxin (50 U) N=105	Placebo N=53	P-value
Investigator Assessment			
Responder	**92 (88%)**	**2 (4%)**	<0.001
Missing	2 (2%)	2 (4%)	
Subject Assessment			
Responder	**78 (74%)**	**5 (9%)**	<0.001
Missing	2 (2%)	2 (4%)	
Composite Assessment			
Responder	**76 (72%)**	**1 (2%)**	<0.001
Missing	2 (2%)	2 (4%)	

Table 7 – Day 30 Response Rates (2+) at Maximum Frown (Study 719)

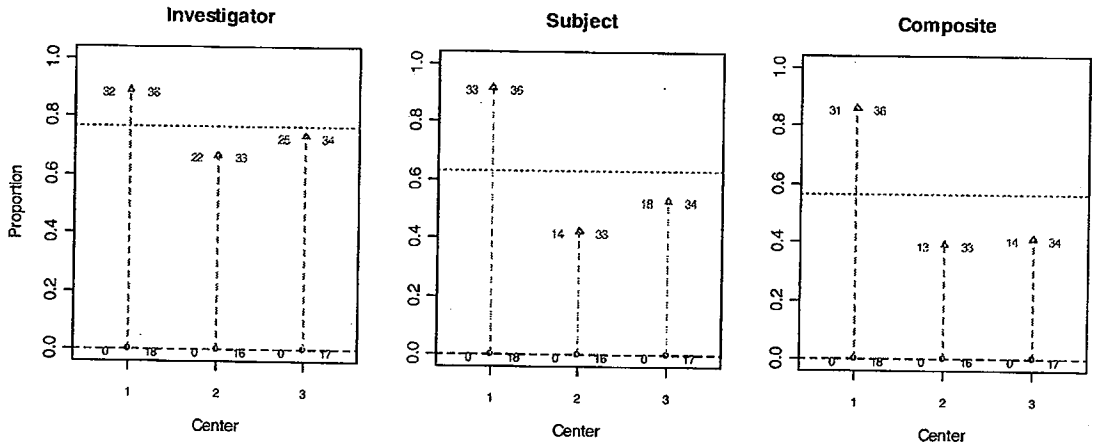
	Reloxin (50 U) N=105	Placebo N=53	P-value
Investigator Assessment			
Responder	**79 (75%)**	**0 (0%)**	<0.001
Missing	2 (2%)	2 (4%)	
Subject Assessment			
Responder	**65 (62%)**	**0 (0%)**	<0.001
Missing	2 (2%)	2 (4%)	
Composite Assessment			
Responder	**58 (55%)**	**0 (0%)**	<0.001
Missing	2 (2%)	2 (4%)	

The applicant handled missing data for the 1+ response rates in two ways: by imputing the average of all non-missing values (both treatments combined, rounded to the nearest integer) for the two treatment groups, and by treating missing values as non-responders. At Day 30, the mean scores at maximum frown for all observed subjects were 1.35 for the investigator assessment and 1.42 for the subject assessment. Thus for the 1+ response rates, missing values in both groups would be imputed as responders under the mean imputation, and as non-responder under the non-response imputation. The applicant did not use any imputation for their presentation of 2+ response rates, presenting results for observed cases only. The applicant did not present per protocol analyses, as the per protocol population was identical to the ITT population. This reviewer's analyses were in agreement with the applicant's analyses. In the reviewer's analyses, subjects with missing values were treated as non-responders.

3.1.1.5 Efficacy by Center

Study 719 used three centers. The 2+ response rates by center are presented in Figure 1. Response rates on the investigator assessment were similar across the centers, but for the subject assessments, the response rate at one of the centers was higher than at the other two centers.

Figure 1 – 2+ Response Rates by Center – Reloxin vs. Placebo (Study 719)

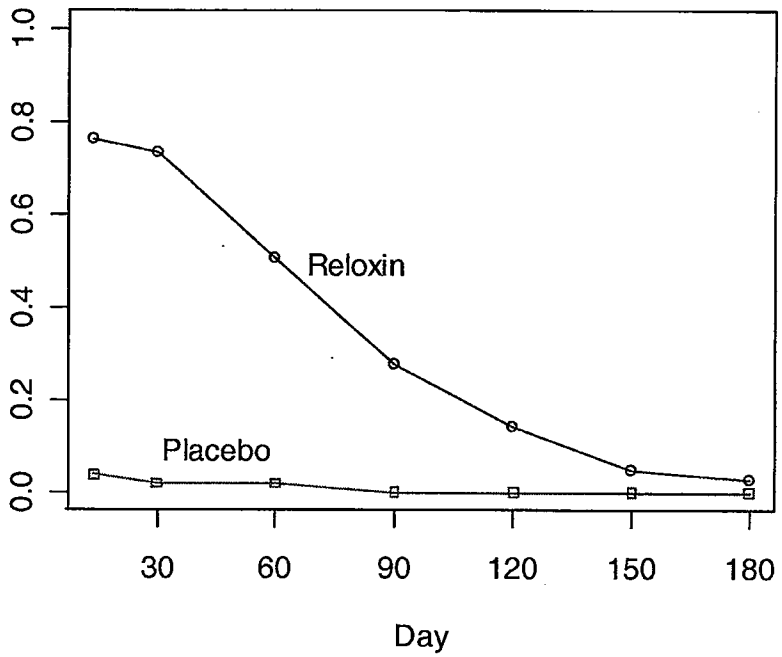


3.1.1.6 Duration of Effect

Subjects in Study 719 were followed every 30 days through Day 180. The proportion of responders decreased over time and very few Reloxin subjects were still in response at the end of the study. By Day 90, about one-third of subjects originally treated with Reloxin were still in 1+ composite response. Figure 2 presents the 1+ composite response rates over time. Although 2+ composite response has advantages when for analyzing the data at a fixed timepoint, as it ensures all responders demonstrated a clinically meaningful response, when tracking over time, having a response definition that does not depend on any previous assessment and treats all subjects with the same score in the same way is preferable. Thus, the figure uses the 1+ composite response as this response rate combines the information from both the subject and investigator assessments, and uses the same criteria for success for all subjects whether they had moderate or severe glabellar lines at baseline.

APPEARS THIS WAY
ON ORIGINAL

Figure 2 – 1+ Composite Response Rate over Time



3.1.2 Study Y-97-52120-085

3.1.2.1 Study Design

Study 085 was a randomized, double-blind, multiple dose study to evaluate the efficacy and safety of Reloxin in the re-treatment of glabellar lines after initial open-label treatments with Reloxin. The study enrolled 311 subjects at 6 U.S. centers. The study enrolled subjects naïve to botulinum toxin treatment at baseline. Subjects were to be age 18 or older with moderate to severe vertical glabellar lines at maximum frown, as assessed by both the subject and investigator. The subject and investigator assessment scales were the same as used in Study 719.

The original study design was to treat all subjects with one or two open-label cycles of Reloxin (Cycles A1 and A2) and then randomize all subjects to treatment with either Reloxin or placebo (Cycle B). However, during the supposedly randomized cycle (B), the sponsor noticed that treatment allocation was highly unbalanced. After an audit, the sponsor discovered all subjects (within the age strata: ≤ 50, >50) were assigned treatments sequentially rather than randomly (Reloxin before placebo). Therefore, after discussions with the Agency, the sponsor modified the protocol to add an additional treatment cycle (Cycle C) for subjects who had continuously received Reloxin, randomizing them to a final cycle of either Reloxin or placebo treatment.

All subjects were treated with Reloxin at baseline (Cycle A1). At Day 120 subjects who were assessed as re-exhibiting moderate or severe glabellar lines at maximum frown on both the investigator's and subject's assessments (henceforth defined as 'relapse') were

retreated with Reloxin if they agreed (Cycle A2). Subjects were then to be randomized to either Reloxin or placebo in Cycle B at the next visit at which they were next eligible for re-treatment (relapse and at least 90 days since last treatment). This would be the first visit with relapse among visits on Day 150, 180, 210, 240, or 270 for subjects who did not have a Cycle A2 treatment, and the first visit with relapse among visits on Day 210, 240, or 270 among subjects who did have a Cycle A2.

However, as noted above, the randomization for Cycle B was botched and subjects were assigned treatments sequentially rather than randomly. The sponsor notified the Agency on April 19, 2006 to request a Type A meeting and presented the information that 171 subjects had been assigned to Treatment 'A' and 26 to 'B', when randomization should have been allocating treatments equally. After meeting with the Agency the sponsor unblinded the data to identify that treatment 'A' was Reloxin and treatment 'B' was placebo. The sponsor then proposed to modify the protocol to add a new randomized cycle. Subjects who received Reloxin in Cycle B were eligible to continue in the study and to be randomized to either Reloxin or placebo in Cycle C when they became eligible for retreatment (relapse and at least 90 days since last treatment). The final treatment allocation in Cycle B was 171 subjects assigned to Reloxin and 84 to placebo. Of the 171 subjects assigned to Reloxin in Cycle B, 142 were eligible and consented to continue on in Cycle C (71 to Reloxin and 71 to placebo). Subjects were randomized without regard to stratum in Cycle C.

Subjects completed a diary card on days 1 through 7 of each treatment cycle to record the onset of treatment. The co-primary efficacy endpoints were 1+ response on the investigator's assessment at maximum frown at Day 30 of Cycle C and the 1+ response on the subject's assessment at maximum frown at Day 30 of Cycle C. At FDA's request, the applicant also conducted an analysis of 1+ composite success from the investigator and subject scores and included analyses for 2+ response. The protocol included a number of secondary endpoints based on independent reviewer's assessments, investigator assessments, subject assessments, and treatment duration assessments at various timepoints.

The Cycle C ITT population was defined as all randomized subjects who received study treatment during Cycle C. The per protocol was to exclude subjects with major protocol violations such as use of non-approved medications, violations of inclusion or exclusion criteria, or missing Day 30 data for Cycle C. No randomized subjects were excluded from the per protocol population so the applicant did not conduct per protocol analyses.

Success rates were analyzed with a Cochran-Mantel-Haenszel test stratified on pooled center and number of treatments in Cycle A (1 or 2). The primary method of handling missing data was to impute the average of all non-missing values (rounded to the nearest integer) for the two treatment groups. Thus, if the average score for the Reloxin and placebo groups combined was less than 1.5, then all missing data (in both arms) would be imputed as responders, and if the average score was greater than 1.5, then all missing data would be imputed as non-responders. As a sensitivity analyses, all missing data would be imputed as non-responders.

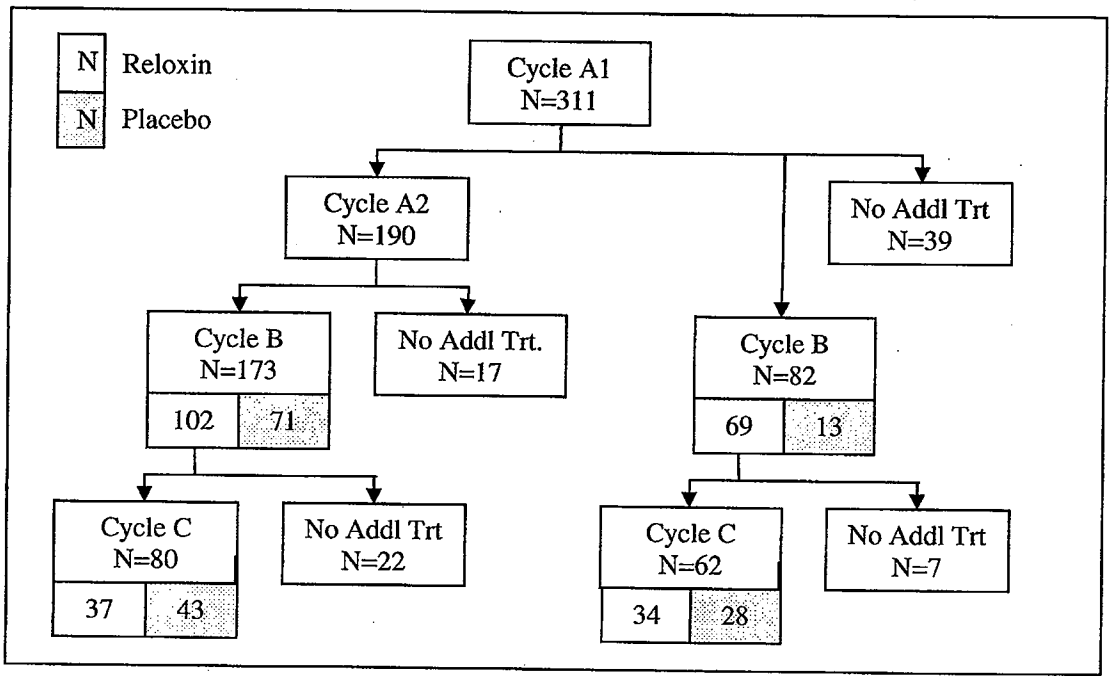
3.1.2.2 Subject Disposition

Study 085 enrolled 311 subjects. By Cycle C, 142 subjects were eligible and treated. Of the 169 subjects who did not enter Cycle C, 52 dropped out during one of the previous cycles and 117 were not eligible or elected not to enter additional cycles. Subjects assigned to placebo in Cycle B were not eligible for Cycle C. Subjects who did not relapse by a certain time were not eligible to enter the next cycle. Table 8 lists the number of subjects treated in each cycle and the reasons for dropout. Figure 3 displays a chart of the number of subjects treated in each cycle. The most common reasons for dropout were loss to follow-up and subject decision.

Table 8 – Subject Disposition (Study 085)

	Reloxin (50 U)	Placebo
Cycle A1	311	
Continued to Cycle A2 or B	272 (87%)	
Discontinued Cycle A1	31 (10%)	
Adverse Event	*1*	
Investigator Decision	*2*	
Subject Decision	*15*	
Non-compliance	*1*	
Lost to Follow-up	*12*	
No Additional Treatment	8 (3%)	
Cycle A2	190	
Continued to Cycle B	173 (91%)	
Discontinued Cycle A2	6 (3%)	
Subject Decision	*2*	
Lost to Follow-up	*4*	
No Additional Treatment	11 (6%)	
Cycle B	171	84
Completed Cycle B and Continued to Cycle C	139 (81%)	--
Discontinued Cycle B but allowed to enter Cycle C	3 (2%)	--
Discontinued Cycle B	9 (5%)	6 (7%)
Subject Decision	*3*	*5*
Lost to Follow-up	*6*	*1*
No Additional Treatment	20 (12%)	78 (93%)
Cycle C	71	71
Attended Day 30 Visit	71 (100%)	71 (100%)
Discontinued Study	1 (1%)	3 (4%)
Subject Decision	*1*	*1*
Lost to Follow-up	--	*2*

Figure 3 – Number of Subjects treated in each Treatment Cycle (Study 085)



3.1.2.3 Baseline Characteristics

At baseline, approximately 85% of the subjects were female. About 80% of the subjects were Caucasian. The mean age was around 46. The demographics at the beginning of Cycle C were similar to baseline and generally balanced across treatment arms, except that the mean age of subjects had dropped to 44. The demographic results are presented in Table 9.

Table 9 – Baseline and Cycle C Demographics (Study 085)

	Baseline	Cycle C	
	Reloxin (50 U) N=311	Reloxin (50 U) N=71	Placebo N=71
Age (years)			
Mean	46.6	44.7	44.7
Range	(21, 74)	(21, 65)	(24, 71)
Gender			
Male	42 (14%)	11 (15%)	9 (13%)
Female	269 (86%)	60 (85%)	62 (87%)
Race			
Caucasian	249 (80%)	57 (80%)	54 (76%)
Hispanic	32 (10%)	6 (8%)	7 (10%)
Afr-Amer	9 (3%)	2 (3%)	4 (6%)
Nat. Amer	4 (1%)	0 (0%)	4 (6%)
Asian	13 (4%)	3 (4%)	2 (3%)
Other	4 (1%)	3 (4%)	0 (0%)

In Study 085, severity at the start of Cycle C was generally balanced across treatment groups, though slightly more severe subjects were randomized to placebo. About 57% of subjects were classified as severe on the investigator and subject assessment scales at the start of Cycle C. See Table 10.

Table 10 – Baseline and Cycle C Assessment at Maximum Frown (Study 085)

	Baseline	Cycle C	
	Reloxin (50 U) N=311	Reloxin (50 U) N=71	Placebo N=71
Investigator			
Moderate	139 (45%)	35 (49%)	25 (35%)
Severe	172 (55%)	36 (51%)	46 (65%)
Subject			
Mild	1 (<1%)	--	--
Moderate	143 (46%)	33 (46%)	30 (42%)
Severe	167 (54%)	38 (54%)	41 (58%)

3.1.2.4 Day 30 of Cycle C Efficacy Results

Reloxin (50 U) was superior to placebo for both protocol-specified primary efficacy endpoints (1+ response at Day 30 on the investigator and subject assessments). Reloxin was also superior to placebo for the investigator and subject 2+ response assessments, as well as both the 1+ and 2+ composite assessments. Efficacy results are presented in Table 11 and Table 12. As all subjects randomized in Cycle C were evaluated at Day 30; no imputation for missing data was necessary. The applicant did not present per protocol analyses, as the per protocol population was identical to the ITT population. This reviewer's analyses were in agreement with the applicant's analyses.

Table 11 – Day 30 of Cycle C Response Rates (1+) at Maximum Frown - ITT (Study 085)

	Reloxin (50 U) N=71	Placebo N=71	P-value
Investigator Assessment			
Responder	**60 (85%)**	**3 (4%)**	<0.001
Missing	0 (0%)	0 (0%)	
Subject Assessment			
Responder	**56 (79%)**	**1 (1%)**	<0.001
Missing	0 (0%)	0 (0%)	
Composite Assessment			
Responder	**54 (76%)**	**0 (0%)**	<0.001
Missing	0 (0%)	0 (0%)	

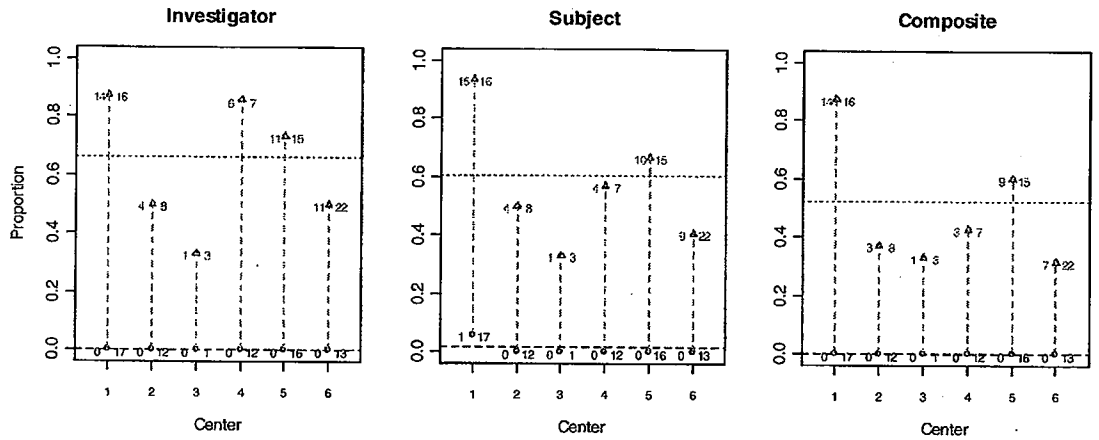
Table 12 – Day 30 of Cycle C Response Rates (2+) at Maximum Frown - ITT (Study 085)

	Reloxin (50 U) N=71	Placebo N=71	P-value
Investigator Assessment			
Responder	**47 (66%)**	**0 (0%)**	<0.001
Missing	0 (0%)	0 (0%)	
Subject Assessment			
Responder	**43 (61%)**	**1 (1%)**	<0.001
Missing	0 (0%)	0 (0%)	
Composite Assessment			
Responder	**37 (52%)**	**0 (0%)**	<0.001
Missing	0 (0%)	0 (0%)	

3.1.2.5 Efficacy by Center

Study 085 used six centers. The 2+ response rates by center are presented in Figure 4. Response rates on the investigator assessment were somewhat variable across the centers. For the subject assessments, the response rate at one of the centers was noticeably higher than at the other five centers.

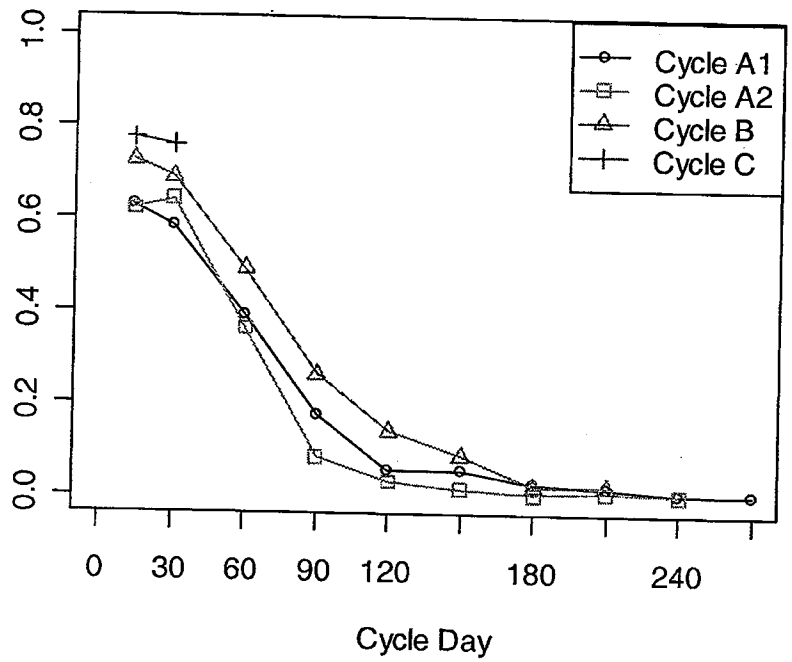
Figure 4 – 2+ Response Rates in Cycle C by Center (Study 085)



3.1.2.6 Efficacy over Time

Subjects in Study 085 were followed in up to 4 treatment cycles. Treatment cycles were a minimum of 90 days apart. In Cycle C, subjects were only evaluated for efficacy through Day 30. The Day 30 1+ composite response rates increased slightly in each subsequent treatment cycle. This may reflect a subject selection process in which subjects who are not responding would be less likely to remain in the study for additional treatments. Response rates (1+ composite) over time for each cycle are presented in Figure 5.

Figure 5 – 1+ Composite Response over Time by Treatment Cycle for Reloxin



3.1.2.7 Randomization Issues

The planned randomization in Cycle B failed to allocate subjects to Reloxin or placebo randomly and instead allocated them sequentially. The sponsor first notified the Agency of a suspected problem in a Type A meeting request dated April 19, 2006 (IND 10673 / 154). In the cover letter of this submission the sponsor stated that it "strongly suspects a serious error in the randomization phase of the study where treatment was allocated sequentially instead of as a true randomization. To date, of the 196 randomized patients, 171 have been allocated to treatment A and 26 have been assigned to treatment B." The Agency and sponsor agreed that to gain any useful information from this study, the blind could be broken, subjects allocated to Reloxin could be re-consented and entered into an additional randomized treatment cycle.

The sponsor commissioned an audit of randomization failure and submitted the report to the IND (Supporting Document Number 167, stamp date July 13, 2006). The key steps that led to the randomization failure were identified as follows:

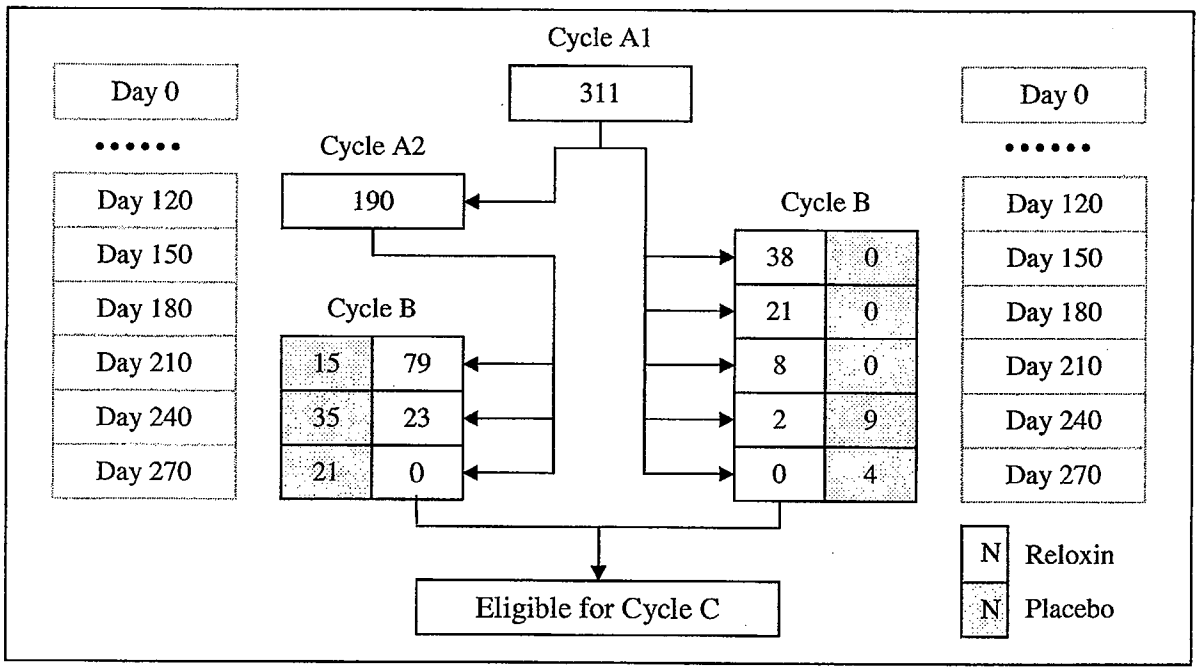
1. The CRO managing the IVRS system requested a 'dummy' randomization list from the sponsor to test the system.
2. The sponsor submitted a dummy randomization list that was 'not downloadable (the format was not correct)' (pg 242 of SDN 167). Thus, testing was done on a dummy list created internally by the CRO, which passed the tests.
3. When the final randomization list was submitted, as testing had not been conducted on randomization lists generated by the sponsor, the teams had not

recognized that the IVRS system could not appropriately handle the format submitted by the sponsor (subject numbers listed by treatment group, rather than in the order they were to be assigned). Thus treatments were assigned in the order they were listed, that is, sequentially rather than randomly.

The Agency has noted since the randomization failure was first identified, that the randomization failure and subsequent needed changes to the protocol may necessitate treating Study 085 as a supportive study rather than a pivotal study. Although the study modification is a serious issue, in the end, the study identified a group of subjects (although the selection process was convoluted), randomized them and followed them after randomization. The key question becomes: are the subjects who entered Cycle C representative of the subjects we would be expecting to evaluate in a trial with open-label lead-in cycles. Figure 6 presents the study days in which subjects received Cycle A2 and Cycle B treatments. Subjects who received Cycle A2 received treatment on Day 120. Subjects who did not receive Cycle A2 and went directly to Cycle B received treatment anywhere between Day 150 and Day 270. The figure shows that of the 171 subjects treated with Reloxin in Cycle B and eligible for Cycle C, most (102/171 or 60%) consistently had relapsed within 120 days of treatment (that is had relapsed by Day 120 and were treated in Cycle A2 and then had relapsed by 90 to 120 days after the Cycle A2 treatment). Most of the remaining subjects eligible for Cycle C maintained response from the initial treatment until the Day 150 or Day 180 visit. The subjects who were not eligible for Cycle C were generally those subjects who had a longer duration response in either Cycle A1 or A2 (approximately 240 days or more for Cycle A1, or 150 days or more for Cycle A2). Of note, the response rates observed in Study 085 are similar to the response rates observed in the other studies, and it does not appear that potential subject selection greatly distorted the results of the study.

APPEARS THIS WAY
ON ORIGINAL

Figure 6 – Study Day of Treatment in Cycles A2 and B (Study 085)



3.1.3 Study A-2006-01

3.1.3.1 Study Design

Study 2006-01 was a randomized, placebo-controlled, double-blind, single dose study to evaluate the efficacy and safety of a higher variable dose of Reloxin in the treatment of glabellar lines. Subjects were treated with a single dose of either Reloxin or placebo. Dosing was based on gender and the investigator's visual assessment of procerus and corrugator muscle mass (+ [small], ++ [medium], +++ [large]), and subjects were assigned to the following doses of Reloxin or comparable volume of placebo.

Table 13 – Dosing Paradigm for Study 2006-01

Muscle mass	Total Dose (U)	
	Female	Male
Small (+)	50	60
Medium (++)	60	70
Large (+++)	70	80

The study enrolled 816 subjects at 27 U.S. centers, 544 treated with Reloxin (50, 60, 70, or 80 U) and 272 treated with placebo. Recruitment was targeted with the goal of enrolling 150 African-American subjects, and 160 African-American subjects were enrolled. Randomization was stratified by race (African-American with Fitzpatrick Skin Types IV, V, or VI, and other). Subjects were not to have had treatment with botulinum toxin within the last 150 days. About 80% of the subjects were botulinum toxin naïve at baseline. Subjects were to be age 18 or older with moderate to severe vertical glabellar lines at maximum frown, as assessed by the subject, investigator, and blinded evaluator.

The blinded evaluator was only responsible for evaluating glabellar severity, while the investigator performed all other investigator responsibilities. The subject and investigator/blinded evaluator assessment scales were the same as used in Study 719.

Subjects completed a diary card on days 1 through 14 to record the onset of treatment. Follow-up clinic visits occurred on Days 14, 30, 60, 90, 120, and 150. The co-primary efficacy endpoints were the duration of efficacy at maximum frown as assessed by the blinded evaluator and the subject. For both assessments, onset was identified by the day on which the subject first recorded a score of 0 or 1. Duration of response was defined by the timepoint at which the score first returned to 2 or 3 on the blinded evaluator or subject assessment, respectively.

The study had 4 rank-ordered key secondary endpoints. These endpoints were 2+ response on the blinded evaluator's assessment at maximum frown on Day 30, 1+ response on the blinded evaluator's assessment at maximum frown on Day 30, duration of 2+ response on the blinded evaluator and subject's assessment at maximum frown, and 1+ response on the blinded evaluator's assessment at rest on Day 30. Based on comments from the Agency dated February 20, 2007 which recommended using 2+ composite response (success on both blinded evaluator and subject assessments), the sponsor modified the first two secondary endpoints to 2+ and 1+ composite response. Also included among the numerous other secondary endpoints were the 1+ response on the subject's assessment at maximum frown on Day 30 and the 1+ response on the investigator's assessment at maximum frown on Day 30, along with assessments at various other timepoints.

The ITT population was defined as all randomized subjects who received study treatment. The per protocol was to exclude subjects with major protocol violations such as use of prohibited or non-approved medications, violations of inclusion or exclusion criteria, missing or out of window study visits, or incomplete study treatment.

Success rates were analyzed with a Cochran-Mantel-Haenszel test stratified on center and race (African-American versus other). Missing data were not imputed (data was analyzed as observed cases only). Duration assessments were analyzed with the Kaplan-Meier method. Duration was measured from the time of response (from subject diary) until the first visit where the subject was no longer a responder.

3.1.3.2 Subject Disposition

Study 2006-01 enrolled 816 subjects. All randomized subjects received their assigned treatment. Only about 1% of subjects in each arm did not attend the Day 30 visit. Dropout by Day 150 in the studies was similar on both arms with about 2% of the subjects dropping out by Day 150. The most common reasons for dropout were loss to follow-up and subject decision. The subject disposition is presented in Table 14.

Table 14 – Subject Disposition (Study 2006-01)

	Reloxin (50 - 80 U)	Placebo
Subjects Randomized	544	272
Attended Day 30 visit	539 (99%)	267 (98%)
Completed Study	534 (98%)	265 (97%)
Discontinuation Reason		
Lost to Follow-Up	7 (1%)	1 (<1%)
Subject Decision	2 (<1%)	6 (2%)
Non-Compliance	1 (<1%)	--

Subjects in Study 2006-01 were assigned to treatment based on their perceived muscle mass and gender. Most subjects had muscle mass recorded as medium or large and thus most females were assigned to 60-70 U of Reloxin and most males were assigned to 70-80 U of Reloxin (or the placebo equivalent). The treatment allocation is presented in Table 15.

Table 15 – Treatment Allocation (Study 2006-01)

	Female N=719	Male N=97
Placebo	238 (33%)	34 (34%)
Reloxin 50 U	22 (3%)	0 (0%)
Reloxin 60 U	277 (39%)	5 (5%)
Reloxin 70 U	181 (25%)	25 (26%)
Reloxin 80 U	1 (<1%)	33 (34%)

3.1.3.3 Baseline Characteristics

Baseline demographics were generally balanced across treatment groups. Approximately 88% of the subjects were female. Almost 70% of the subjects were Caucasian. The study made a special effort to enroll African-American subjects and approximately 20% of subjects were African-American. The mean age was around 49. The demographic results are presented in Table 16.

APPEARS THIS WAY
ON ORIGINAL

Table 16 – Baseline Demographics (Study 2006-01)

	Reloxin (50-80 U) N=544	Placebo N=272
Age (years)		
Mean	48.7	49.2
Range	(20, 78)	(23, 80)
Gender		
Male	63 (12%)	34 (13%)
Female	481 (88%)	238 (88%)
Race		
Caucasian	364 (67%)	191 (70%)
Hispanic	57 (10%)	19 (7%)
African-American	106 (19%)	54 (20%)
Native American	1 (<1%)	2 (<1%)
Asian	8 (1%)	3 (1%)
Other	8 (1%)	3 (1%)

In Study 2006-01, baseline severity was generally balanced across treatment groups. About 66% of subjects were classified as severe on the investigator assessment and about 56% of subjects were classified as severe on the subject assessment scale. See Table 17. One subject (84.017) had a blinded evaluator assessment at maximum frown of 0 at baseline, although the same blinded evaluator's assessment at *rest* was a score of 2. In addition, the enrolling investigator assessed the subject with a score of 2 at both maximum frown and rest. It is possible that the blinded evaluator miscoded the maximum frown response, though the data are presented and analyzed as recorded in the CRF.

Table 17 – Baseline Assessment at Maximum Frown (Study 2006-01)

	Reloxin (50-80 U) N=544	Placebo N=272
Blinded Evaluator		
None	1 (<1%)	--
Moderate	193 (35%)	81 (30%)
Severe	350 (64%)	191 (70%)
Subject		
Moderate	249 (46%)	113 (42%)
Severe	295 (54%)	159 (58%)

3.1.3.4 Day 30 Efficacy Results

Reloxin (50-80 U) was superior to placebo for each of the Day 30 response rates. Efficacy results are presented in Table 18 and Table 19. Day 30 response rates were secondary endpoints in Study 2006-01. The applicant analyzed the Day 30 response rates using observed cases only, without using any imputation. The percentages in the following tables are out of the number of randomized subjects (not out of the number of observed subjects). Thus, these tables treat missing values as non-responders. Except for

the handling of missing data, the reviewer's analyses are in agreement with the applicant's analyses.

Table 18 – Day 30 Response Rates (1+) at Maximum Frown (Study 2006-01)

	Reloxin (50-80 U) N=544	Placebo N=272	P-value
Blinded Evaluator Assessment			
Responder	**455 (84%)**	**9 (3%)**	<0.001
Missing	7 (1%)	5 (2%)	
Subject Assessment			
Responder	**469 (86%)**	**12 (4%)**	<0.001
Missing	6 (1%)	5 (2%)	
Composite Assessment			
Responder	**428 (79%)**	**6 (2%)**	<0.001
Missing	7 (1%)	0 (0%)	

Note: All percentages are out of the number of randomized subjects.

Table 19 – Day 30 Response Rates (2+) at Maximum Frown (Study 2006-01)

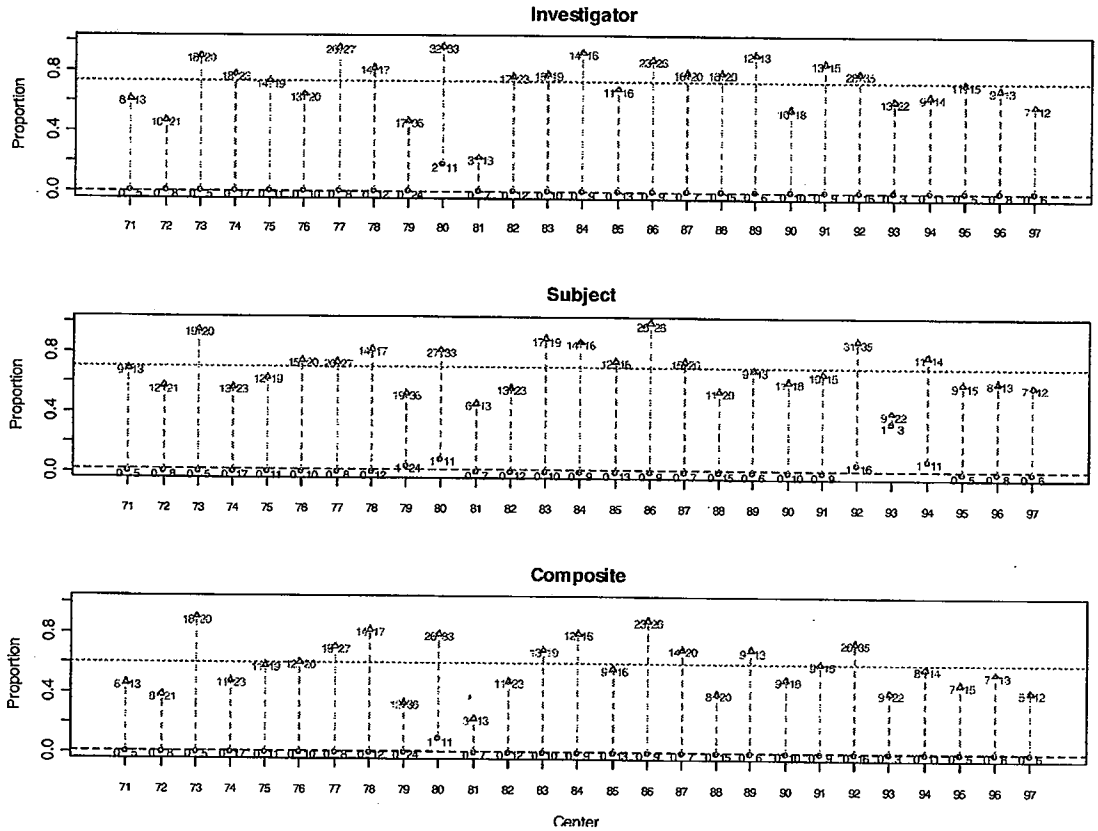
	Reloxin (50-80 U) N=544	Placebo N=272	P-value
Blinded Evaluator Assessment			
Responder	**398 (73%)**	**2 (<1%)**	<0.001
Missing	7 (1%)	5 (2%)	
Subject Assessment			
Responder	**379 (70%)**	**5 (2%)**	<0.001
Missing	6 (1%)	5 (2%)	
Composite Assessment			
Responder	**319 (59%)**	**1 (<1%)**	<0.001
Missing	7 (1%)	5 (2%)	

Note: All percentages are out of the number of randomized subjects.

3.1.3.5 Efficacy Results by Center

Study 2006-01 used 27 centers. The 2+ response rates by center are presented in Figure 7. Response rates on the investigator and subject assessments were somewhat variable, but generally similar across the centers. No center appears to dominate the efficacy results. Center was not considered in the randomization, and thus, although the randomization was 2:1, many centers did not end up with a comparable allocation. For example, Center 93 had 22 Reloxin and 3 placebo subjects, while Center 74 had 23 Reloxin and 17 placebo subjects.

Figure 7 – 2+ Response Rates by Center (Study 2006-01)

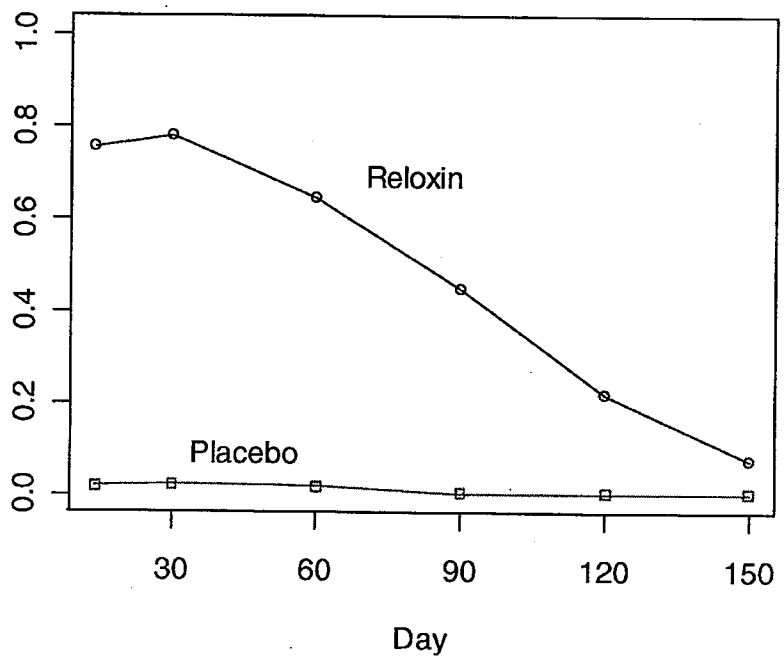


3.1.3.6 Efficacy over Time

Subjects in Study 2006-01 were followed every 30 days through Day 150. The proportion of responders decreased over time and very few Reloxin subjects were still in response at the end of the study. By Day 90, about 45% of subjects originally treated with Reloxin were still in 1+ composite response. Figure 8 presents the 1+ composite response rates over time.

APPEARS THIS WAY
ON ORIGINAL

Figure 8 – 1+ Composite Response Rates over Time (Study 2006-01)



Duration of 1+ response on the blinded evaluator and subject assessments were designated as the primary efficacy endpoints in Study 2006-01. For both the blinded evaluator and subject assessments of duration, the response start date was determined from the subject's diary card from the first 14 days. Duration is defined as the number of days for a responder to re-exhibit an assessment of 2 or 3 on either the blinded evaluator or subject assessment following onset of treatment response. The median duration of response on Reloxin was 109 days for the blinded evaluator assessment and 107 days for the subject assessment. The median duration on placebo was 0 days for both assessments. Duration results are presented in Table 20 and Table 21.

APPEARS THIS WAY
ON ORIGINAL

Table 20 – Kaplan-Meier Estimates of Duration of Blinded Evaluator Response (Study 2006-01)

	Reloxin N = 544	Placebo N = 272
Number of subjects who responded to treatment	511 (94%)	31 (11%)
Number (%) of subjects who became non-responders during the study observation period	426 (83%)	30 (97%)
Number of subjects censored	85 (17%)	1 (3%)
Median duration of response to treatment (days)	109	0
p-value (Wilcoxon test)	<0.001	
Kaplan Meier estimate for the probability of being a responder for:		
14 days	0.858	0.066
30 days	0.831	0.022
60 days	0.742	0.015
90 days	0.603	0.007
120 days	0.383	0.004
150 days	0.155	0.004

Source: Table 7, pg 90, file a-2006-01-study-report-body.pdf.

Table 21 – Kaplan-Meier Estimates of Duration of Subject Response (Study 2006-01)

	Reloxin N = 544	Placebo N = 272
Number of subjects who responded to treatment	511 (94%)	31 (11%)
Number (%) of subjects who became non-responders during the study observation period	426 (83%)	27 (87%)
Number of subjects censored	85 (17%)	4 (13%)
Median duration of response to treatment (days)	107	0
p-value (Wilcoxon test)	<0.001	
Kaplan Meier estimate for the probability of being a responder for:		
14 days	0.846	0.077
30 days	0.822	0.055
60 days	0.735	0.033
90 days	0.606	0.026
120 days	0.387	0.018
150 days	0.157	0.015

Source: Table 8, pg 92, file a-2006-01-study-report-body.pdf.

3.1.3.7 Variable Dose

In the protocol for Study 2006-01 the applicant used the following rationale for using the variable dosing scheme: "as published references establish, the administered dose of botulinum Type A toxin should be varied within a narrow range based on the muscle mass of the frontalis, procerus and corrugator muscles. In general clinicians recommend administration of slightly higher delivered doses in males as compared to females." (pg 13, file a-2006-01-protocol-or-amendment.pdf). The applicant has not done any dose-ranging work comparing doses within muscle mass or gender categories to demonstrate that gender or muscle mass are important factors in dose selection. Without this comparative information, it is impossible to tell whether the higher doses actually provide an efficacy benefit. While it is noticeable in other studies that efficacy in males is lower than in females, that does not necessarily mean that a higher dose would lead to greater efficacy. The only available data on the response rates of males at different doses is to use response rates from different studies and any comparisons could be subject to many sources of bias. None of the other studies besides 2006-01 collected information on muscle mass, so it is impossible to tell if subjects in previous studies with larger muscle mass had lower efficacy than subjects with smaller muscle mass. If the applicant wishes to pursue variable dosing, then the applicant should conduct studies demonstrating that gender and muscle mass are important factors and that increasing the dose based on those factors increases efficacy without sacrificing safety.

Response rates (1+) by unit dose or corresponding placebo volume are presented in Table 22. The 1+ response rates are presented so that any placebo response trends can be assessed. The 2+ response yields very little placebo response and thus any trends in placebo response would be indiscernible. From this table, the main finding is that most placebo response is contributed by subjects with the smallest muscle mass (0.4 mL placebo dose for females and 0.5 mL placebo dose for males), although relatively few subjects had the smallest muscle mass. It is not possible to tell from this table whether subjects receiving the higher doses actually received any efficacy benefit from the higher doses.

Table 22 – 1+ Response Rates by Unit Dose or Corresponding Placebo Volume (Study 2006-01)

	Female		Male	
	Reloxin	Placebo	Reloxin	Placebo
Blinded Evaluator				
50 U / 0.4 mL	21/22 (91%)	1/11 (9%)	--	--
60 U / 0.5 mL	247/276 (89%)	6/139 (4%)	5/5 (100%)	1/3 (33%)
70 U / 0.6 mL	146/177 (82%)	1/84 (1%)	16/25 (64%)	0/4 (0%)
80 U / 0.7 mL	1/1 (100%)	--	19/32 (59%)	0/26 (0%)
Subject				
50 U / 0.4 mL	21/22 (91%)	1/11 (9%)	--	--
60 U / 0.5 mL	245/276 (89%)	10/139 (7%)	5/5 (100%)	1/3 (33%)
70 U / 0.6 mL	155/178 (87%)	0/84 (0%)	18/25 (72%)	0/4 (0%)
80 U / 0.7 mL	1/1 (100%)	--	24/32 (75%)	0/26 (0%)

3.1.4 CAMR Product Studies

The final Phase 3 study that will be evaluated in this review is Study 718. However, Study 718 used an earlier version of the bulk active substance referred to as CAMR product. The later studies used bulk active substance referred to as IBL product. The CAMR and IBL products have been noted to have some analytical differences. Study 718 will be reviewed as a supportive study. The applicant also conducted a bridging study comparing subjects treated with CAMR and IBL product. The bridging study, Study 096, will be briefly reviewed to provide rationale for considering Study 718 as a supportive efficacy and safety study.

3.1.4.1 Study Y-97-52120-096

3.1.4.1.1 Study Design

Study 096 was a randomized, double-blind Phase 2 study designed primarily to assess the relative safety of CAMR and IBL Reloxin. The single-center study treated 50 subjects with 50 U of CAMR Reloxin and 50 subjects with 50 U of IBL Reloxin and followed them for 30 days. Evaluating the 1+ response on the investigator and subject assessments at maximum frown on Day 30 were secondary objectives of the study. Subjects were to be naïve to botulinum toxin treatment. Subjects were to be age 18 or older with moderate to severe vertical glabellar lines at maximum frown, as assessed by both the subject and investigator. The subject and investigator assessment scales were the same as used in Study 719.

The protocol defined primarily descriptive analyses, with no formal efficacy hypothesis testing. No non-inferiority margins for evaluating the relative efficacy of the two formulations were specified. The study reported point estimates and 95% confidence intervals (normal approximation) for the difference in 1+ response rates.

During the study, 4 IBL and 2 CAMR subjects discontinued before Day 30. In the IBL group, 3 discontinued due to subject request and 1 was lost to follow-up. In the CAMR group, 1 discontinued due to subject request and 1 was lost to follow-up. Subjects with missing data were treated as non-responders.

3.1.4.1.2 Day 30 Efficacy Results

In this small study, the IBL response rates were similar to slightly lower than the CAMR response rates. The 1+ response rates are presented in Table 23. Although this study was not designed to formally assess bioequivalence, the results indicate that the two products have similar response rates and that Study 718 (which used the CAMR product) may provide useful supporting efficacy information.

Table 23 – Day 30 Response Rates (1+) at Maximum Frown (Study 096)

	IBL (50 U) N=50	CAMR (50 U) N=50	95% Conf. Int. (IBL – CAMR)
Investigator Assessment			
Responder	**40 (80%)**	**44 (88%)**	(-22%, 6%)
Missing	4 (8%)	2 (4%)	
Subject Assessment			
Responder	**40 (80%)**	**40 (80%)**	(-16%, 16%)
Missing	4 (8%)	2 (4%)	
Composite Assessment			
Responder	**38 (76%)**	**39 (78%)**	(-18%, 14%)
Missing	4 (8%)	2 (4%)	

3.1.4.2 Study Y-97-52120-718

3.1.4.2.1 Study Design

Study 718 was a randomized, placebo-controlled, double-blind, single dose study to evaluate the efficacy and safety of Reloxin (CAMR) in the treatment of glabellar lines. This study was completed using CAMR product, prior to the switch to the to-be-marketed IBL product. Subjects were treated with a single dose of either Reloxin (50 U-CAMR) or placebo and followed for 150 days. The study enrolled 300 subjects at 5 U.S. centers, 200 treated with Reloxin and 100 treated with placebo. Randomization was stratified by age (≤ 50, >50). Subjects were to be naïve to botulinum toxin treatment. Subjects were to be age 18 or older with moderate to severe vertical glabellar lines at maximum frown, as assessed by both the subject and investigator. The subject and investigator assessment scales were the same as used in Study 719. In addition to the investigator's and subject's assessment, each subject's photographs were evaluated by 3 independent reviewers using the investigator scale.

Subjects completed a diary card on days 1 through 7 to record the onset of treatment. Follow-up clinic visits occurred on Days 14, 30, 60, 90, 120, and 150. The co-primary efficacy endpoints were the investigator's assessment at maximum frown at Day 30 and the subject's assessment at maximum frown at Day 30. For each scale a responder was defined as having a score of 2 or 3 at baseline and a score of 0 or 1 at Day 30. The protocol included a number of secondary endpoints based on independent reviewer's assessments, investigator assessments, subject assessments, and treatment duration assessments at various timepoints.

The ITT population was defined as all randomized subjects who received study treatment. The per protocol was to exclude subjects with major protocol violations such as violations of inclusion or exclusion criteria, procedural, etc. No subjects were found to have had any major protocol violations and the per protocol population was identical to the ITT populations. However, the applicant reported that one center (05) was 'relatively inexperienced at clinical research' and failed to maintain all original source documentation, did not complete study documentation in real time, and did not complete

all study activities in the correct sequence. Thus the applicant defined an MITT population excluding this center.

Success rates were analyzed with a Cochran-Mantel-Haenszel test stratified on age (≤ 50, > 50). Duration of response was analyzed with a log-rank test stratified on age (≤ 50, > 50). The primary method of handling missing data was to impute the average of all non-missing values (rounded to the nearest integer) for the two treatment groups separately. Thus, if the average score for the Reloxin or placebo group was less than 1.5, then all missing data in that arm would be imputed as responders, and if the average score was greater than 1.5, then all missing data in that arm would be imputed as non-responders. As a sensitivity analyses, all missing data would be imputed as non-responders.

3.1.4.2.2 Subject Disposition

Study 718 enrolled 300 subjects. All randomized subjects received their assigned treatment. About 5% of subjects in each arm did not attend the Day 30 visit. Dropout by Day 150 in the studies was similar on both arms with about 6% of the subjects dropping out by Day 150. The most common reasons for dropout were loss to follow-up and subject decision. The subject disposition is presented in Table 24.

Table 24 – Subject Disposition (Study 718)

	Reloxin (50 U)	Placebo
Subjects Randomized	200	100
Attended Day 30 visit	192 (96%)	93 (93%)
Completed Study	190 (95%)	92 (92%)
Discontinuation Reason		
Lost to Follow-Up	6 (3%)	3 (3%)
Subject Decision	2 (1%)	3 (3%)
Lack of Efficacy	1 (<1%)	--
Non-Compliance	1 (<1%)	2 (2%)

3.1.4.2.3 Baseline Characteristics

Baseline demographics were generally balanced across treatment groups. Approximately 87% of the subjects were female. About 75% of the subjects were Caucasian and 19% were Hispanic. The mean age was around 44. The demographic results are presented in Table 25.

APPEARS THIS WAY
ON ORIGINAL

Table 25 – Baseline Demographics (Study 718)

	Reloxin (50 U) N=200	Placebo N=100
Age (years)		
Mean	44.7	43.2
Range	(21, 71)	(24, 66)
Gender		
Male	28 (14%)	12 (12%)
Female	172 (86%)	88 (88%)
Race		
Caucasian	149 (75%)	76 (76%)
Hispanic	37 (19%)	18 (18%)
African-American	5 (3%)	5 (5%)
Native American	1 (1%)	0 (0%)
Asian	4 (2%)	0 (0%)
Other	4 (2%)	1 (1%)

In Study 718, baseline severity was balanced across treatment groups. About 65% of subjects were classified as severe on both the investigator and subject assessment scales. See Table 26.

Table 26 – Baseline Assessment at Maximum Frown (Study 718)

	Reloxin (50 U) N=200	Placebo N=100
Investigator		
Moderate	64 (32%)	31 (31%)
Severe	136 (68%)	69 (69%)
Subject		
Moderate	74 (37%)	37 (37%)
Severe	126 (63%)	63 (63%)

3.1.4.2.4 Day 30 Efficacy Results

Reloxin (50 U) was superior to placebo for both protocol-specified primary efficacy endpoints (1+ response at Day 30 on the investigator and subject assessments). Reloxin was also superior to placebo for the investigator and subject 2+ response assessments, as well as both the 1+ and 2+ composite assessments. Efficacy results are presented in Table 27 and Table 28.

Table 27 – Day 30 Response Rates (1+) at Maximum Frown (Study 718)

	Reloxin (50 U) N=200	Placebo N=100	P-value
Investigator Assessment			
Responder	**171 (86%)**	**0 (0%)**	<0.001
Missing	8 (4%)	7 (7%)	
Subject Assessment			
Responder	**163 (82%)**	**2 (2%)**	<0.001
Missing	8 (4%)	7 (7%)	
Composite Assessment			
Responder	**152 (76%)**	**0 (0%)**	<0.001
Missing	8 (4%)	7 (7%)	

Table 28 – Day 30 Response Rates (2+) at Maximum Frown (Study 718)

	Reloxin (50 U) N=200	Placebo N=100	P-value
Investigator Assessment			
Responder	**157 (79%)**	**0 (0%)**	<0.001
Missing	8 (4%)	7 (7%)	
Subject Assessment			
Responder	**136 (68%)**	**0 (0%)**	<0.001
Missing	8 (4%)	7 (7%)	
Composite Assessment			
Responder	**120 (60%)**	**0 (0%)**	<0.001
Missing	8 (4%)	7 (7%)	

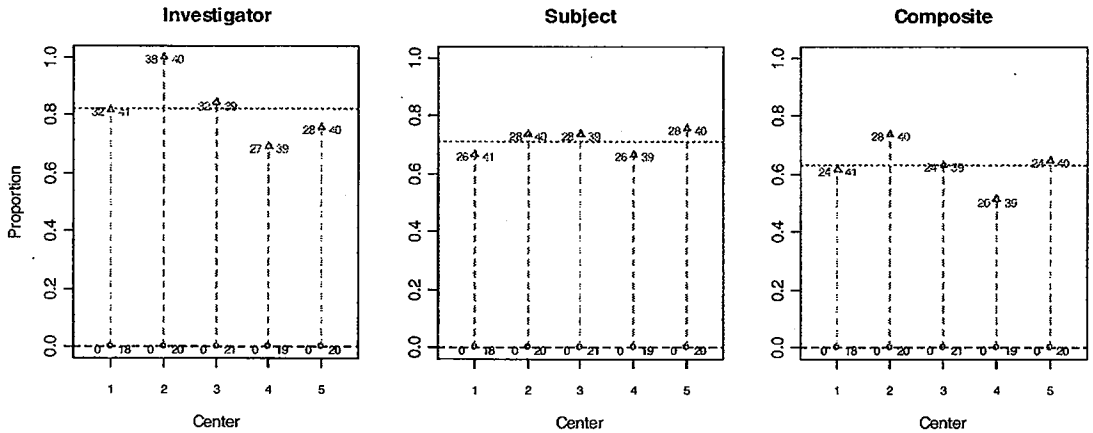
The applicant handled missing data for the 1+ response rates in two ways: by imputing the average of all non-missing values (both treatments separately rather than combined, rounded to the nearest integer) for the two treatment groups, and by treating missing values as non-responders. At Day 30, the mean investigator scores at maximum frown were 0.67 for Reloxin and 2.73 for placebo. The mean subject scores at maximum frown were 0.77 for Reloxin and 2.54 for placebo. Thus for the 1+ response rates, missing values in the Reloxin group would be imputed as responders and missing values in the placebo group would be imputed as non-responders under the mean imputation. The applicant did not conduct analyses of 2+ response rates. The applicant did not present per protocol analyses, as the per protocol population was identical to the ITT population. The reviewer analyses are in agreement with the applicant's analyses, except that the reviewer analyses treat all subjects with missing data as non-responders, rather than treating missing Reloxin subjects as responders and missing placebo subjects as non-responders.

3.1.4.2.5 Efficacy by Center

Study 718 used five centers. The 2+ response rates by center are presented in Figure 9. Response rates on the investigator and subject assessments were similar across the centers. The applicant reported that Center 5 was 'relatively inexperienced at clinical research' and failed to maintain all original source documentation, did not complete study

documentation in real time, and did not complete all study activities in the correct sequence. However, the response rates at Center 5 are similar to the other centers.

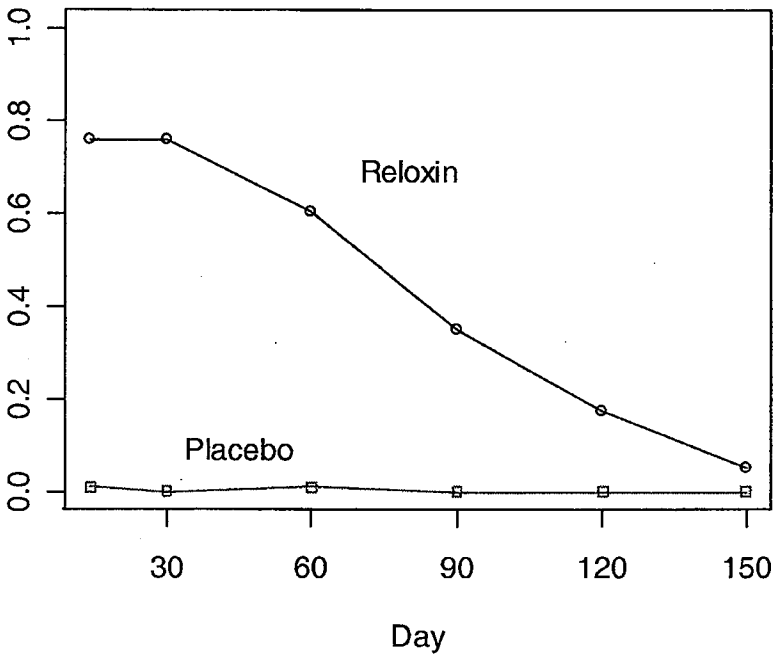
Figure 9 – 2+ Response Rates by Center (Study 718)



3.1.4.2.6 Efficacy over Time

Subjects in Study 718 were followed every 30 days through Day 150. The proportion of responders decreased over time and very few Reloxin subjects were still in response at the end of the study. By Day 90, about one-third of subjects originally treated with Reloxin were still in 1+ composite response. Figure 10 presents the 1+ composite response rates over time.

Figure 10 – 1+ Composite Response Rates over Time (Study 718)



3.2 Evaluation of Safety

In all studies, Reloxin had a higher rate of adverse events than placebo. Much of the difference appears to be due to eye disorders such as ptosis and injection site reactions. (See Table 29 through Table 34.) Ptosis was observed in about 2% of subjects. In Study 2006-01, ptosis was observed at a higher rate in subjects receiving 70 U of Reloxin (8/206 or 3.8%) than in subjects receiving 60 U of Reloxin (4/281 or 1.4%). (See Table 33.) Additionally, 2 of the affected 70 U subjects had ptosis in both eyes and one subject had severe ptosis. There were too few subjects receiving 50 U or 80 U of Reloxin in Study 2006-01 to make comparisons with these dose levels.

Table 29 – Adverse Events Occurring in at least 2 Reloxin Subjects (≥ 2%) (Study 719)

	Reloxin (50U) N=105	Placebo N=53
Any Adverse Event	49 (47%)	21 (40%)
Eye Disorders		
Blepharospasm	1 (1%)	2 (4%)
Eyelid Ptosis	3 (3%)	0 (0%)
Administration Site Conditions		
Injection Site Reaction	8 (8%)	0 (0%)
Injection Site Pain	4 (4%)	2 (4%)
Other		
Nasopharyngitis	12 (11%)	6 (11%)
Headache	10 (10%)	4 (8%)
Vomiting	3 (3%)	1 (2%)
Influenza	2 (2%)	2 (4%)
Nausea	2 (2%)	1 (2%)

APPEARS THIS WAY
ON ORIGINAL

Table 30 - Adverse Events Occurring in at least 5 Reloxin Subjects (≥ 2%) during the Study (Study 085)

	Reloxin (50U) N=311	Placebo N=155
Any Adverse Event	196 (63%)	43 (28%)
Eye Disorders		
Eyelid Ptosis	6 (2%)	--
Asthenopia	5 (2%)	--
Dry Eye	5 (2%)	--
Administration Site Conditions		
Injection Site Reaction	5 (2%)	--
Injection Site Pain	11 (4%)	1 (<1%)
Injection Site Bruising	10 (3%)	6 (4%)
Other		
Headache	38 (12%)	4 (3%)
Nasopharyngitis	31 (10%)	6 (4%)
Sinusitis	13 (4%)	1 (<1%)
Hypertension	10 (3%)	1 (<1%)
Influenza	9 (3%)	1 (<1%)
Acne	8 (3%)	--
Bronchitis	6 (2%)	1 (<1%)
Upper Respiratory Tract Infection	6 (2%)	--
Back Pain	6 (2%)	2 (1%)
Depression	6 (2%)	1 (<1%)
Nausea	6 (2%)	--
Rash	6 (2%)	--
Seasonal Allergy	5 (2%)	2 (1%)

Note: Cumulative events are from all cycles. Subjects may have received multiple Reloxin treatments. Subjects may have received Reloxin or placebo in different cycles.

Table 31 - Adverse Events Occurring in at least 2 Reloxin Subjects (≥ 2%) during Cycle C (Study 085)

	Reloxin (50U) N=71	Placebo N=71
Any Adverse Event	27 (38%)	21 (30%)
Headache	3 (4%)	2 (3%)
Nasopharyngitis	9 (13%)	4 (6%)
Bronchitis	1 (1%)	3 (4%)
Nausea	2 (3%)	--
Diarrhea	2 (3%)	--

Table 32 - Adverse Events Occurring in at least 6 Reloxin Subjects (≥ 1%) during the Study (Study 2006-01)

	Reloxin (50-80U) N=544	Placebo N=272
Any Adverse Event	168 (31%)	75 (28%)
Eyelid Ptosis	13 (2%)	1 (<1%)
Procedural Pain	6 (1%)	--
Headache	19 (3%)	8 (3%)
Nasopharyngitis	15 (3%)	6 (2%)
Upper Respiratory Tract Infection	10 (2%)	4 (1%)
Sinusitis	6 (1%)	3 (1%)

Table 33 – Subjects with Ptosis (Study 2006-01)

Subject	Eye Affected	Onset Day	Duration	Severity
Placebo (N = 272) – 1 (>1%)				
83.013 (F)	Left	0	1 day	Mild
Reloxin 50 U (N = 22) – 1 (5%)				
76.018 (F)	Right	10	35 days	Mild
Reloxin 60 U (N = 281) – 4 (1%)				
76.005 (F)	Right	3	47 days	Mild
76.030 (F)	Right	17	23 days	Mild
79.063 (F)	Left	9	Continuing	Moderate
97.012 (F)	Right	8	45 days	Mild
Reloxin 70 U (N = 206) – 8 (4%)				
72.020 (F)	Left	1	8 days	Mild
72.020 (F)	Right	1	8 days	Mild
75.007 (F)	Right	18	2 days	Severe
75.007 (F)	Right	20	16 days	Mild
75.030 (F)	Right	18	12 days	Mild
78.024 (F)	Right	16	126 days	Mild
82.013 (F)	Not Indicated	14	82 days	Mild
85.028 (F)	Right	12	83 days	Moderate
86.003 (F)	Left	22	21 days	Mild
86.003 (F)	Right	22	21 days	Mild
92.051 (F)	Left	25	1 day	Mild
Reloxin 80 U (N = 34) No reports of ptosis.				

Table 34 - Adverse Events Occurring in at least 4 Reloxin Subjects (≥ 2%) during the Study (Study 718)

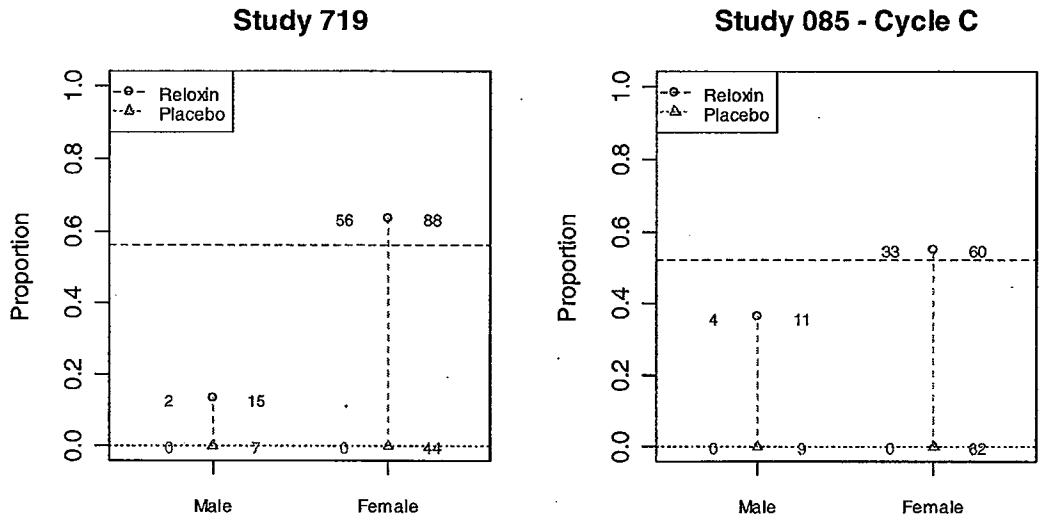
	Reloxin (50 U) N=200	Placebo N=100
Any Adverse Event	104 (52%)	43 (43%)
Eye Disorders		
Eyelid Edema	6 (3%)	--
Eyelid Ptosis	2 (1%)	--
Administration Site Conditions		
Injection Site Pain	7 (4%)	1 (1%)
Injection Site Reaction	4 (2%)	2 (2%)
Injection Site Swelling	4 (2%)	1 (1%)
Injection Site Discomfort	4 (2%)	--
Other		
Headache	22 (11%)	9 (9%)
Nasopharyngitis	16 (8%)	4 (4%)
Upper Respiratory Tract Infection	8 (4%)	5 (5%)
Sinusitis	4 (2%)	2 (2%)
Nasal Congestion	4 (2%)	--
Blood Urine Present	6 (3%)	1 (1%)

4 Findings in Special/Subgroup Populations

4.1 Gender, Race, and Age

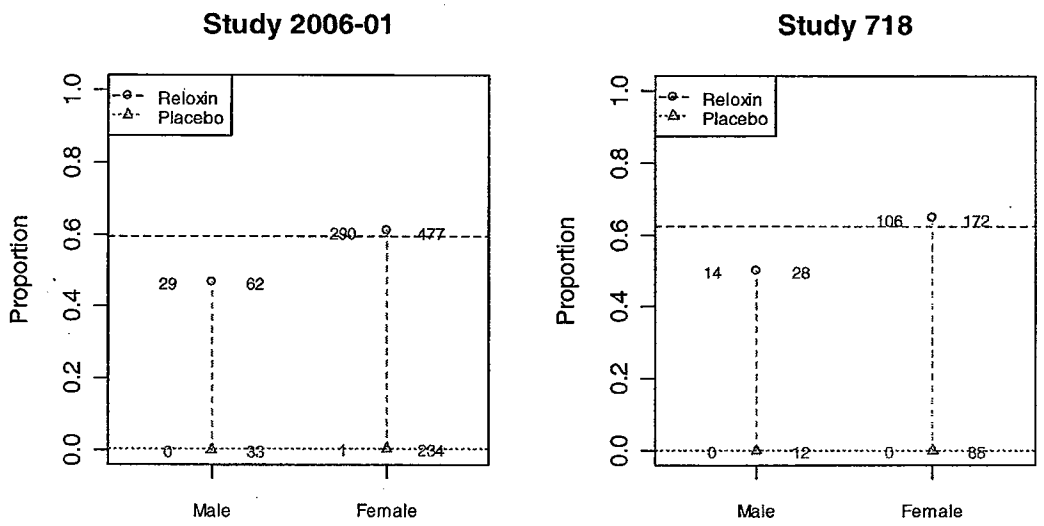
Several subgroup result patterns were observed across all studies: females had higher response rates than males, non-Caucasians had higher response rates than Caucasians, and younger subjects (≤ 50) had higher response rates than older subjects (> 50). See Figure 11 through Figure 16. Part of the applicant's rationale for pursuing a variable dosing scheme in Study 2006-01 was an attempt to address the finding that males often have lower response rates than females. The only study that conducted dose ranging was Study 717 (which is not reviewed here), and the application did not include raw data or by-gender efficacy tables for Study 717. Thus only cross-study comparisons about the impact on dosing and gender can be made. Recognizing the limited utility of cross-study comparisons, it appears that the within-gender results of Study 2006-01 (variable dosing) are similar to the results from the fixed dose studies 085 and 718. Only Study 719 has substantially lower response rates for males than the other studies. In all cases, the number of male subjects is small. However, the applicant has not provided compelling evidence that male subjects benefit substantially from higher doses of Reloxin.

Figure 11 – 2+ Composite Response Rates at Day 30 by Gender (Studies 719 and 085)



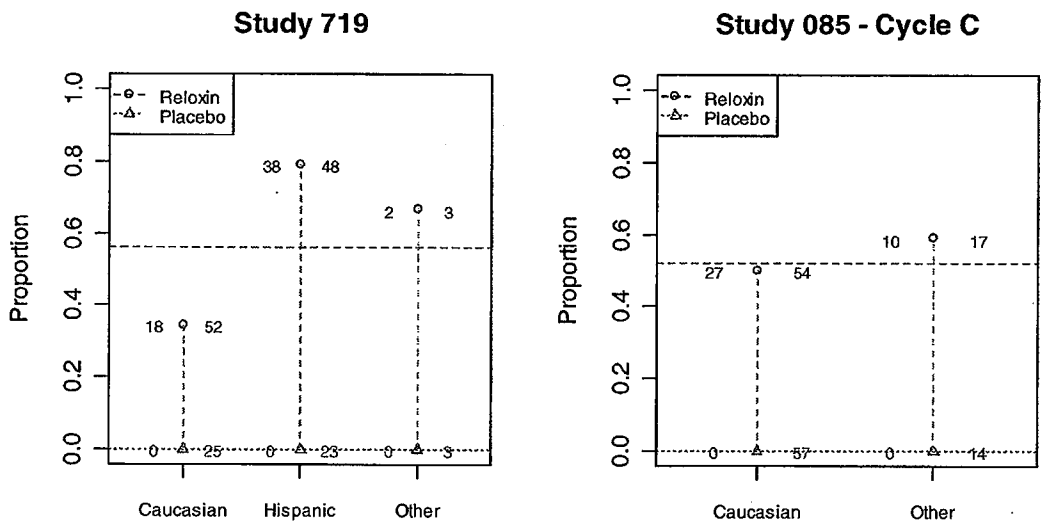
Note: Numbers represent the number of successes (on the left) and the number of subjects (on the right).

Figure 12 - 2+ Composite Response Rates at Day 30 by Gender (Studies 2006-01 and 718)



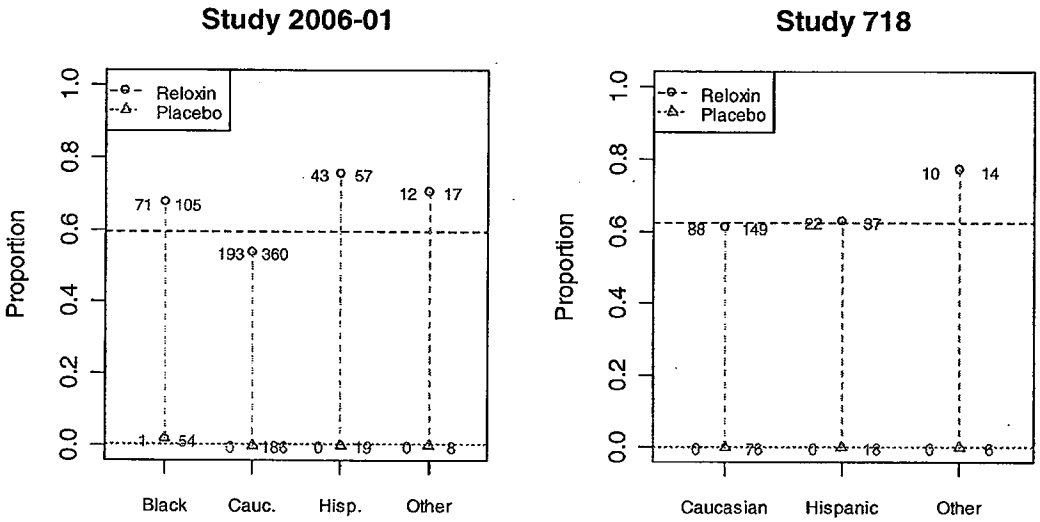
Note: Numbers represent the number of successes (on the left) and the number of subjects (on the right).

Figure 13 - 2+ Composite Response Rates at Day 30 by Race (Studies 719 and 085)



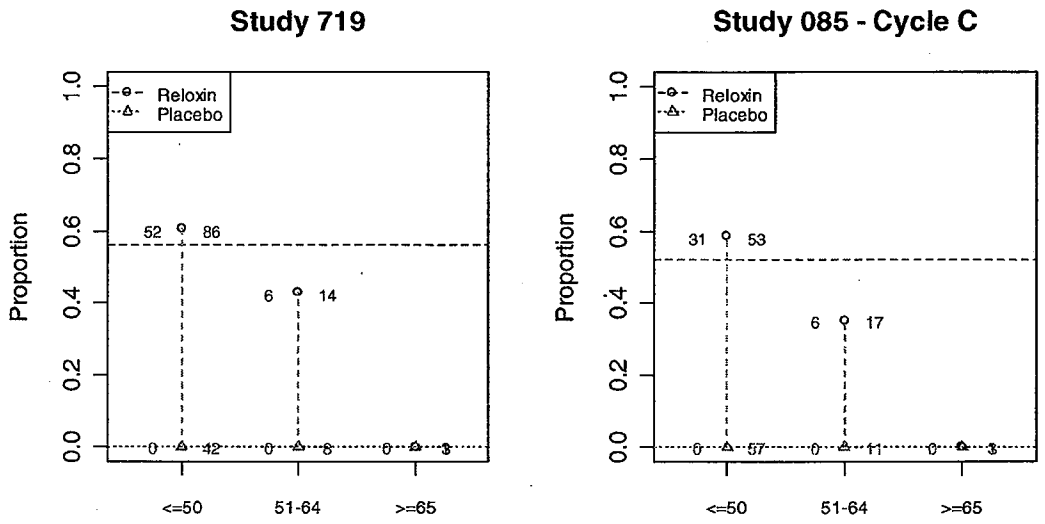
Note: Numbers represent the number of successes (on the left) and the number of subjects (on the right).

Figure 14 - 2+ Composite Response Rates at Day 30 by Race (Studies 2006-01 and 718)



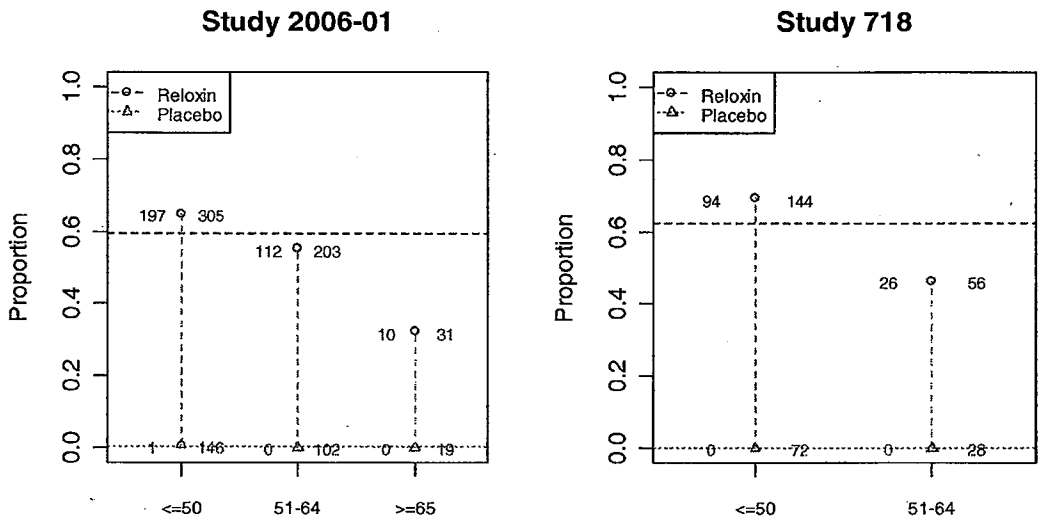
Note: Numbers represent the number of successes (on the left) and the number of subjects (on the right).

Figure 15 - 2+ Composite Response Rates at Day 30 by Age Group (Studies 719 and 085)



Note: Numbers represent the number of successes (on the left) and the number of subjects (on the right).

Figure 16 - 2+ Composite Response Rates at Day 30 by Age Group (Studies 2006-01 and 718)



Note: Numbers represent the number of successes (on the left) and the number of subjects (on the right).

4.2 Other Special/Subgroup Populations

Not applicable.

5 Summary and Conclusions

5.1 Statistical Issues and Collective Evidence

The applicant has conducted four Phase 3 efficacy and safety studies of Reloxin, each conducted under slightly different conditions

- 718 – single dose, 50 U vs. placebo, CAMR product
- 719 – single dose, 50 U vs. placebo, IBL product
- 085 – multiple dose, 50 U vs. placebo, IBL product (with redesign due to randomization problems)
- 2006-01 – single dose, 50-80 U vs. placebo, IBL product

In addition the applicant has conducted a dose-ranging study (Study 717 using 20, 50, or 75 U of CAMR product), a small comparability study between the CAMR and IBL product (Study 096), and two long-term safety studies (720 and 732).

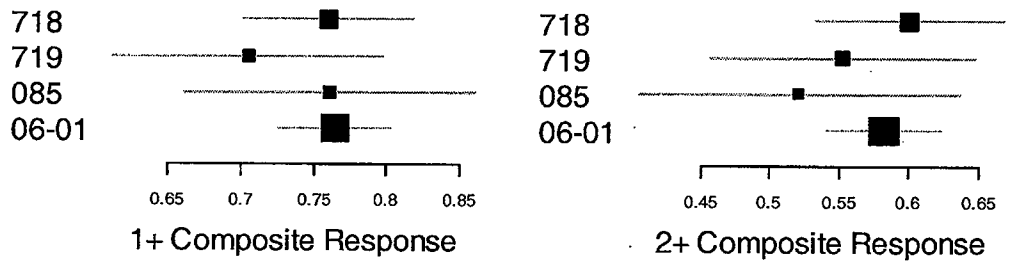
5.1.1 Fixed versus Variable Dosing

The only safety information on doses greater than 50 U of IBL product comes from Study 2006-01. All of the submitted data from the long-term safety studies used the 50 U dose of Reloxin. Although Study 2006-01 was the largest efficacy study with over 800 subjects, it was a single-dose study, and no repeat-dose information on doses higher than 50 U has been submitted. Although it has been consistently observed across studies that males tend to have lower response rates than females, the applicant has not provided convincing evidence that higher doses of Reloxin lead to higher response rates in males. No within-study comparative data have been submitted, and other studies that treated males with 50 U of Reloxin (085, 718) had similar response rates in males as Study 2006-01. Higher levels of ptosis were observed with 70 U than 60 U of Reloxin in Study 2006-01 (3.8% vs. 1.4%). In addition, the applicant also noted that the 50 U dose was originally selected because the 75 U dose in Study 717 also had higher levels of ptosis than the 50 U dose. Thus although the applicant has demonstrated that the variable dosing scheme is an effective dose, the applicant has not presented convincing evidence that (1) there is an efficacy benefit to the higher doses over the fixed dose, (2) that any potential efficacy benefit justifies increases in safety events such as ptosis, or that (3) long-term safety of the higher doses has been adequately evaluated.

5.1.2 Efficacy of 50 U Dose

All four Phase 3 studies had relatively similar efficacy results, even though the study designs and products and doses differed from study to study. Table 35 presents 1+ and 2+ composite response rates. Figure 17 presents a forest plot of the confidence intervals for the treatment effects (1+ and 2+ composite response) in the four studies. Treatment effect estimates (Reloxin – Placebo) for 1+ composite response varied from 70% to 76% and the treatment effect estimates for 2+ composite response varied from 52% to 60%. Thus, the different 'products'/doses evaluated (50 U IBL, 50 U CAMR, 50-80 U IBL) in the single dose studies (719, 718, 2006-01) and the repeat dose study (085) all led to similar response rates and treatment effects.

Figure 17 – Treatment Effect Forest Plots - 95% Confidence Intervals for Difference in Response Rates (Reloxin - Placebo)



Note: Box sizes are determined by the standard errors. Studies with larger sample sizes have larger boxes.

Table 35 – Composite Response Rates in Phase 3 Studies

Study	Reloxin	Placebo	P-value
1+ Composite Response			
718	152/200 (76%)	0/100 (0%)	<0.001
719	76/105 (72%)	1/53 (2%)	<0.001
085	54/71 (76%)	0/71 (0%)	<0.001
2006-01	428/544 (79%)	6/272 (2%)	<0.001
2+ Composite Response			
718	120/200 (60%)	0/100 (0%)	<0.001
719	58/105 (55%)	0/53 (0%)	<0.001
085	37/71 (52%)	0/71 (0%)	<0.001
2006-01	319/544 (59%)	1/272 (<1%)	<0.001

5.1.3 Statistical Issues

Study 719 demonstrated that 50 U of Reloxin is superior to placebo in the treatment of glabellar lines. This review did not identify any significant statistical concerns regarding the design or conduct of Study 719 that would impact the efficacy conclusions.

In Study 085 the failed randomization is a significant concern. However, because the goal of the study was to assess efficacy after a few lead-in cycles of Reloxin, adding an additional cycle may have impacted the selection of subjects who entered the final randomized cycle, but otherwise the study was able to evaluate a group of randomized subjects and assess efficacy. The efficacy results in Study 085 are consistent with the results from other Reloxin studies.

Study 2006-01 was able to demonstrate that a variable dose of Reloxin is superior to placebo. Most subjects in Study 2006-01 received either 60 or 70 U of Reloxin. Although most subjects received a higher dose of Reloxin, the response rates in Study 2006-01 were fairly comparable to those observed in other studies. Because the dose response curve does not appear to be very steep within the range of 50 to 80 U, Study 2006-01 provides additional support that Reloxin is efficacious, even though relatively

few subjects in the study received 50 U, and it cannot be directly inferred that efficacy at a higher dose implies efficacy at a lower dose.

Study 718 demonstrated the efficacy of CAMR Reloxin. The small comparability Study 096 provides some additional support that IBL product is not substantially less effective than CAMR product, although this study was not designed to establish equivalence between the products. Thus Study 718 may also provide supportive evidence of efficacy for Reloxin.

5.2 Conclusions and Recommendations

The applicant has demonstrated the efficacy of Reloxin (50 U) in the treatment of glabellar lines relative to placebo in two studies (719 and 085). This review did not identify any significant statistical concerns regarding the design or conduct of Study 719 that would impact the efficacy conclusions. Study 085 had a problem with its initial randomization, which required the study to be modified with an additional treatment cycle. Subjects who had received Reloxin in all previous treatment cycles were eligible to be randomized into the final added cycle. Although the study design was modified, all modifications were completed before subjects were randomized into the final treatment cycle, and thus the results are interpretable.

Two additional studies also provide supportive information regarding the efficacy of Reloxin. Study 2006-01 demonstrated the efficacy of a variable dose of Reloxin (50-80 U) with dosing based on gender and the investigator's visual assessment of procerus and corrugator muscle mass. However, the applicant has not provided sufficient evidence that the higher doses provide greater efficacy (no comparative dose-ranging information has been provided), and a higher rate of ptosis was observed with 70 U of Reloxin in female subjects. Study 718 demonstrated that an earlier version of Reloxin using bulk active substance from a different source (referred to as CAMR product) also was superior to placebo.

The co-primary efficacy endpoints in Studies 719 and 085 were the investigator's assessment at maximum frown at Day 30 and the subject's assessment at maximum frown at Day 30. Responders on the investigator and subject scales had scores of 'none' or 'mild' (defined here as 1+ response). In Study 085 the endpoints were assessed at Day 30 in Cycle C. Subjects eligible for Cycle C had previous received 2 or 3 treatments with Reloxin (Cycles A1, A2 (certain subjects only), and B). In addition, the studies also demonstrated efficacy on response definitions of interest to the Agency: 1+ composite response (none or mild on *both* the investigator and subject assessments), and 2+ composite response (none or mild with at least 2 grades reduction from baseline on both the investigator and subject assessments). The 2+ composite response rates are presented in Table 36. The totality of evidence from the clinical studies supports the claim that Reloxin (50 U) has demonstrated efficacy in the treatment of glabellar lines.

Table 36 – 2+ Composite Response Rates in Phase 3 Studies

Study	Dose/Product	Reloxin	Placebo	P-value
719	50 U	58/105 (55%)	0/53 (0%)	<0.001
085	50 U	37/71 (52%)	0/71 (0%)	<0.001
2006-01	50 – 80 U	319/544 (59%)	1/272 (<1%)	<0.001
718	50 U (CAMR)	120/200 (60%)	0/100 (0%)	<0.001

Signatures/Distribution List

Primary Statistical Reviewer: Kathleen Fritsch, PhD 11/21/08
Date: 11/21/2008

Statistical Team Leader: Mohamed Alosh, PhD 11/21/08

cc:
DDDP/Walker
DDDP/Alosh
DDDP/Cook
DDDP/White
OBIO/Patrician
DBIII/Wilson
DBIII/Alosh
DBIII/Fritsch

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

BLA Number: 125286 **Applicant:** Ipsen **Stamp Date:** 3/14/2008

Drug Name: Reloxin **NDA/BLA Type:**

On initial overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	√			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	√			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).	√			
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).	√			

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? Yes

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	√			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	√			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.			√	
Appropriate references for novel statistical methodology (if present) are included.			√	
Safety data organized to permit analyses across clinical trials in the NDA/BLA.	√			
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.	√			

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Reviewing Statistician	4/28/08 Date
Supervisor/Team Leader	4/28/08 Date

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207