



Food and Drug Administration  
Center for Biologics Evaluation and Research  
Office of Biostatistics and Epidemiology  
Division of Biostatistics

# STATISTICAL REVIEW AND EVALUATION BLA

**BLA/Supplement Number:** 125287/0 /25 and later amendments

**Product Name:** Berinert

**Indication(s):** Treatment of acute abdominal or facial attacks of hereditary angioedema (HAE) in adult and adolescent patients

**Applicant:** CSL Behring

**Date(s):** March 7th, 2008 ( CBER receipt date of original submission)  
April 10<sup>th</sup>, 2009 (CBER receipt date of STN 125287/0 /25)

**Review Priority:** Standard

**Statistical Branch:** Therapeutics Evaluation Branch

**Primary Statistical Reviewer:** Xue Lin, Ph.D.

**Concurring Reviewer (1):** Jessica (Jeongsook) Kim PhD., team leader, Therapeutics Evaluation Branch, (HFM-219)

**Concurring Reviewer (2):** Ghanshyam Gupta, PhD., Chief, Therapeutics Evaluation Branch, (HFM-219)

**Medical Office/Division:** OBRR/DH

**Clinical Reviewer(s):** Leland Ross Pierce, MD (HFM-380)

**Project Manager:** Nanette Cagungun (HFM-380)

# Table of Contents

<b>1. EXECUTIVE SUMMARY .....</b>	<b>3</b>
<b>2. INTRODUCTION .....</b>	<b>3</b>
2.1 OVERVIEW.....	3
2.2 DATA SOURCES .....	4
<b>3. STATISTICAL EVALUATION .....</b>	<b>4</b>
3.1 EVALUATION OF EFFICACY .....	5
3.2 EVALUATION OF SAFETY .....	8
3.3 SUBGROUP ANALYSIS.....	8
<b>4. SUMMARY AND CONCLUSIONS .....</b>	<b>9</b>
<b>APPENDIX .....</b>	<b>9</b>
<b>DISTRIBUTION LIST.....</b>	<b>11</b>

## **1. EXECUTIVE SUMMARY**

This BLA application was submitted by CSL Behring (CSLB) for Berinert in treating acute abdominal or facial attacks of hereditary angioedema (HAE) in adult and adolescent patients. This submission includes one pivotal study which is a randomized, three-arm, placebo-controlled, double-blind, dose-finding, multinational study. The primary objective of the study is to show that Berinert shortens the time to onset of relief of symptoms of abdominal or facial HAE attack compared to placebo. The primary efficacy endpoint is time between start of study medication administration and onset of relief of symptoms from evaluated (abdominal or facial) attack determined by subject's assessment.

This reviewer confirmed the sponsor's p-value of 0.014 for the primary efficacy analysis based on the revised data they provided (the sponsor referred it as "robustness analysis"). The FDA clinical reviewer performed a masked review of the submitted source documents. Based on the clinical reviewer's data, this reviewer derived primary efficacy data which were different from the sponsor's. At the clinical reviewer's suggestion, this reviewer conducted three analyses and all three had a one-sided p-value less than 0.0249 (0.0053, 0.0016 and 0.0004 respectively using the Wilcoxon rank sum test with normal approximation). The p-values for both the secondary endpoint analyses are less than 0.1 (0.000084 and 0.033 respectively using Fisher's exact test and Wilcoxon rank sum test). According to the study protocol, the one-sided p-value for the final analysis needs to be less than 0.0249 to win the primary efficacy endpoint. Also, based on agreement between FDA and the sponsor, at least one one-sided p-value for the final analysis of the two secondary endpoints needs to be less than 0.1 to win the trial. The study results meet both criteria.

## **2. INTRODUCTION**

### **2.1 Overview**

This BLA application was submitted by CSL Behring for Berinert in treating acute abdominal or facial attacks of hereditary angioedema (HAE) in adult and adolescent patients. Berinert is a plasma derived concentrate of C1 Esterase Inhibitor (Human) for intravenous use only. This submission includes one pivotal study which is a randomized, three-arm, placebo-controlled, double-blind, dose-finding, multinational study. The primary objective of the study is to show that Berinert shortens the time to onset of relief of symptoms of abdominal or facial HAE attack compared to placebo. The secondary objectives are to compare the efficacy of two different dosing schemes of Berinert in

abdominal or facial HAE attack and to compare the safety profiles of initial three treatment groups.

The original BLA application was submitted on March 7, 2008 and the statistical analysis plan was reviewed by Dr. Chinying Wang. On December 5, 2008 FDA issued a complete response (CR) letter. The major clinical issues raised in the CR letter were on concomitant medications. According to the protocol, if potentially confounding concomitant medications were taken in a certain time window, the primary endpoint value would be rated as poor/failure (24 hours). FDA pointed out that the sponsor's list of potentially confounding concomitant medications does not contain all medications of analgesic or anti-emetic property as it should be. FDA requested the sponsor to update their list and redo their analysis. Also, FDA would like the sponsor to extend the time window for counting potentially confounding concomitant medications from beginning of study medication to 5 hours prior to that. A third major issue was that the administration time of many concomitant medications was not recorded. In the CR letter, FDA asked the sponsor to provide source documents including hospital records for review.

On April 10<sup>th</sup>, 2009, FDA received CSLB's response to the CR letter. In their response, CSLB insisted that non-narcotic pain medication should not be included in the prohibited medication list and cited outside doctor's letters. They submitted the analysis FDA required but denoted them as "robustness analysis". In the new analysis they submitted, the sponsor updated the prohibited concomitant medication list to include additional prohibited concomitant medications and extended time window starting at 5 hours prior of study medication. Also they revised their primary endpoint data based on source medical documents.

On May 29, 2009, FDA sent a Request for Information (RFI) to CSLB. At FDA's request, the sponsor provided one complete data set for validation of the primary efficacy analysis (called "robustness analysis" by the sponsor). This statistical review is on this most up-to-date primary efficacy data. CSLB also submitted concomitant medication source documents and copies of hospital medication records for FDA's review.

## **2.2 Data Sources**

- Related materials in supplement 25 and 31, including data files such as `crl_q6cn.xpt`, `adcm.xpt`, `adhaesym.xpt`, `adhaebas.xpt`, `advomit.xpt`.
- The FDA clinical reviewer (Dr. LeLand Ross Pierce) provided Berinert-Imputation-Concom-Meds-X-Rescue (5).xpt for evaluation of the primary efficacy endpoint based on his review of source medical documents.

## **3. STATISTICAL EVALUATION**

### 3.1 Evaluation of Efficacy

#### Study Design and Endpoints

The pivotal study is a randomized, three-arm, placebo-controlled, double-blind, dose-finding, multinational study. The primary efficacy endpoint is time between start of study medication administration and onset of relief of symptoms from evaluated (abdominal or facial) attack determined by subject's assessment.

The secondary efficacy endpoints include:

2. Proportion of subjects with worsened intensity of clinical HAE symptoms between 2 and 4 hours after start of study medication administration compared to baseline, for at least one of the HAE symptoms present at baseline
3. Number of vomiting episodes within 4 hours after start of study treatment

Subjects with C1-INH deficiency with an acute moderate to severe abdominal or facial attack were eligible for participation. It was required that at least 70% of treatment group had abdominal attacks. Only one single abdominal or facial attack of a subject was evaluated for this study, even when multiple attacks were present at baseline. The observation period began with enrollment into the study and treatment of the abdominal or facial attack.

The study started with three arms: high dose Berinert (20 units/kg body weight), low dose Berinert (10 units/kg body weight) and placebo. At least twenty-five evaluable subjects (per-protocol population) were planned to be enrolled in each arm. It was planned that when 15 subjects in each arm were evaluable, an interim analysis would be performed and the DSMB would recommend the sponsor either to go on as planned, or to increase the sample size, or to stop for futility.

The interim analysis was performed as planned. The low dose arm was ceased for futility and the recalculated sample size was 100 subjects for both high dose arm and placebo. Despite the FDA clinical reviewer's recommendation to follow the recalculated sample size, the sponsor decided to enroll a maximum of 42 subjects per treatment group due to financial considerations. This statistical review is based on the available data.

After the study had begun and some subjects had already received the treatment, the sponsor proposed a second interim analysis. The objective of the second interim analysis was to make decisions on whether to stop for futility, or go on as planned, or to stop for overwhelming efficacy. The sponsor proposed spending 0.0001 alpha at this interim analysis and 0.0249 at the final analysis. This interim analysis plan was reviewed by Dr. Chinying Wang.

The second interim analysis was conducted as planned when there were 25 evaluable subjects in both high dose and placebo groups. The trial didn't stop and continued till full enrollment.

### **Patient Disposition, Demographic and Baseline Characteristics**

One hundred and twenty six subjects were enrolled in the study: 42 in the placebo arm, 40 in the low dose arm, 43 in the high dose arm and one subject was not randomized. Low dose arm was ceased after the first interim analysis. The 85 subjects in the placebo and high dose arms composed the population for the primary efficacy analysis.

The following table shows the demographic and baseline characteristics of the placebo and high dose arms.

Category	Placebo (n=42)	High Dose Berinert (n=43)
<b>Age</b>		
3-12 yrs	2 (4.8%)	1 (2.3%)
12-17 yrs	3 (7.1%)	4 (9.3%)
17-65 yrs	37(88.1%)	35 (81.4%)
65+ yrs	0	3 (7.0%)
<b>Gender</b>		
Male	14 (33.3%)	13 (30.2%)
Female	28 (66.7%)	30 (69.8%)
<b>Race</b>		
Caucasian	37 (88.1%)	38 (88.4%)
Other	5 (11.9%)	5 (11.6%)
<b>Baseline Attack Type</b>		
Abdominal	33 (78.6%)	34 (79.1%)
Facial	8 (19.1%)	9 (21.0%)
Larynx	1 (2.4%)	0 (0%)

### **Statistical Methodologies**

In the final protocol the sponsor proposed the two-sample Wilcoxon test to compare the primary endpoint: time between start of study medication administration (ToS) and onset of relief of symptoms (ToSRel) in the high dose Berinert arm and the placebo arm. One-sided significance level of 0.025 was used.

If subjects received rescue medicine or analgesics/anti-emetics concomitant medications before ToSRel was reached or ToSRel-ToS > 24 hours or ToSRel cannot be determined because of missing values, the primary endpoint was set to 24 hours. Otherwise, it was equal to ToSRel- ToS.

The two secondary efficacy endpoints were analyzed with one-sided Fisher's exact test and one-sided two-sample Wilcoxon test, respectively.

The sponsor agreed that to claim the trial successful, besides winning the primary endpoint, they need to show that the one-sided p-value of at least one of the two secondary endpoints was less than 0.1.

To justify that type I error was under control, the sponsor cited Lan and Trost (*Proceedings of the Biopharmaceutical Section*, 1997; pp. 48-51) and Chen (*Statistics in Medicine*, 2004, 1023-10382). Applying Lan and Trost, the re-estimated sample size was to be 100 per arm. However, the sponsor enrolled only about 40 subjects in each arm due to their financial considerations. There was no statistical justification regarding how much influence this deviation had on the type I error rate. The second interim analysis was proposed after the study was already initiated. Without prospective plan to control overall type I error taking all the interim analyses into account, it is hard to justify that type I error is still under control in this complicated study design.

## **Results and Conclusions**

### **Sponsor's results**

The sponsor's one-sided p-value for the primary efficacy analysis was 0.014. The one-sided p-values for the 2 secondary endpoints are 0.0014 (increased intensity of clinical HAE symptoms) and 0.033 (number of vomiting episodes), respectively. Both are less than 0.1.

### **FDA Statistical Reviewer's findings**

This reviewer confirmed the sponsor's p-value of 0.014 based on the primary efficacy data the sponsor provided.

Among the 85 subjects in the ITT population for the primary efficacy analysis, 42 subjects took potentially confounding concomitant medications and from CRF it could not be determined if the medicine was taken in the time window (5 hours before study medication till start relief of symptom) in which case a poor/failure value (24 hours) should be assigned for the primary efficacy endpoint. On case by case bases, the sponsor reviewed hospital source records, the investigator query form response together with CRF data and determined whether the medication was taken in the time window. However, the FDA clinical reviewer found that there are cases in which the hospital source records were not available or were inadequate or the CRF records were inconsistent or conflicts with the investigator query form response. So he suggested three analyses as follows:

(a) FDA will assign "poor/failure" 24 hour values for the primary efficacy endpoint for all cases in which potentially confounding concomitant medications (other than androgens prescribed to be taken on a regular basis, but not prn androgens and aspirin in the dose of 81 mg or less) were marked on the CRF as "ongoing" and where the sponsor has not submitted any medication source documents, or where the source documents and CRF data were judged ambiguous as to whether such medications were taken between -5 hours and ToSRel

(b) FDA will assign “poor/failure” 24 hour values for the primary efficacy endpoint for all cases in which potentially confounding concomitant medications (other than androgens prescribed to be taken on a regular basis, but not prn androgens and aspirin in the dose of 81 mg or less) were marked on the CRF as “ongoing” and where provided source documents suggest a reasonable possibility that such medications may have been taken during the time window of interest (i.e., where such source documents are judged less definitive than a clear hospital record showing no such medications being taken during the time window of interest)

(R2) FDA will accept the statements on the investigator query form as factual if it was received by the sponsor within 180 days of the date of randomization, and the FDA analysis will accept the CRF box checked “ongoing” and assume the medication (other than androgens prescribed to be taken on a regular basis, but not prn androgens and aspirin in the dose of 81 mg or less) was administered between 5 hours prior to ToS and ToSRel if the investigator query form was received > 6 months from the date of randomization (or if other source document indicate potentially confounding medications (including, but not limited to PROHMED dataset medications) were plausibly taken during the time window of interest). Review of concomitant medications taken during -5 hours prior to randomization and ToSRel revealed that clonazepam and mirtazepine, both benzodiazepines, should have been included on the PROHMED list but were not. 24 hours will not be imputed for subjects solely because they took a non-narcotic analgesic between -5 hours and ToSRel.

The one-sided p-values for the three analyses are 0.0053 for the analysis (a), 0.0016 for the analysis (b) and 0.0004 for the analysis (R2). The p-values were obtained based on one-sided two sample Wilcoxon rank sum test with normal approximation. All three p-values are less than 0.0249, so the sponsor wins the primary efficacy endpoint based on any of the three analyses. Kaplan-Meier curves were investigated for the three analyses (See the Appendix).

Note that this reviewer’s results are different from the sponsor’s in one of the two secondary endpoints. For increased intensity of clinical HAE symptoms, the sponsor’s one-side p-value is 0.0014 and this reviewer’s is 0.000084. For the number of vomiting episodes, this reviewer result is the same as the sponsor’s (one-sided p-value=0.033). However, the conclusion will not be affected by the discrepancy.

### **3.2 Evaluation of Safety**

There are no specific safety analyses that are crucial for product approval.

### **3.3 Subgroup Analysis**

Primary efficacy analysis was performed in two baseline HAE attack type subgroups: Abdominal and facial attacks. The p-values for analysis (a) using one-sided two sample Wilcoxon rank sum test with normal approximation are 0.0066 for the abdominal attack



subgroup and 0.301 for the facial attack subgroup. Note that the sample size in the facial attack subgroup is small with only 17 subjects (8 in placebo and 9 in high dose Berinert).

#### **4. SUMMARY AND CONCLUSIONS**

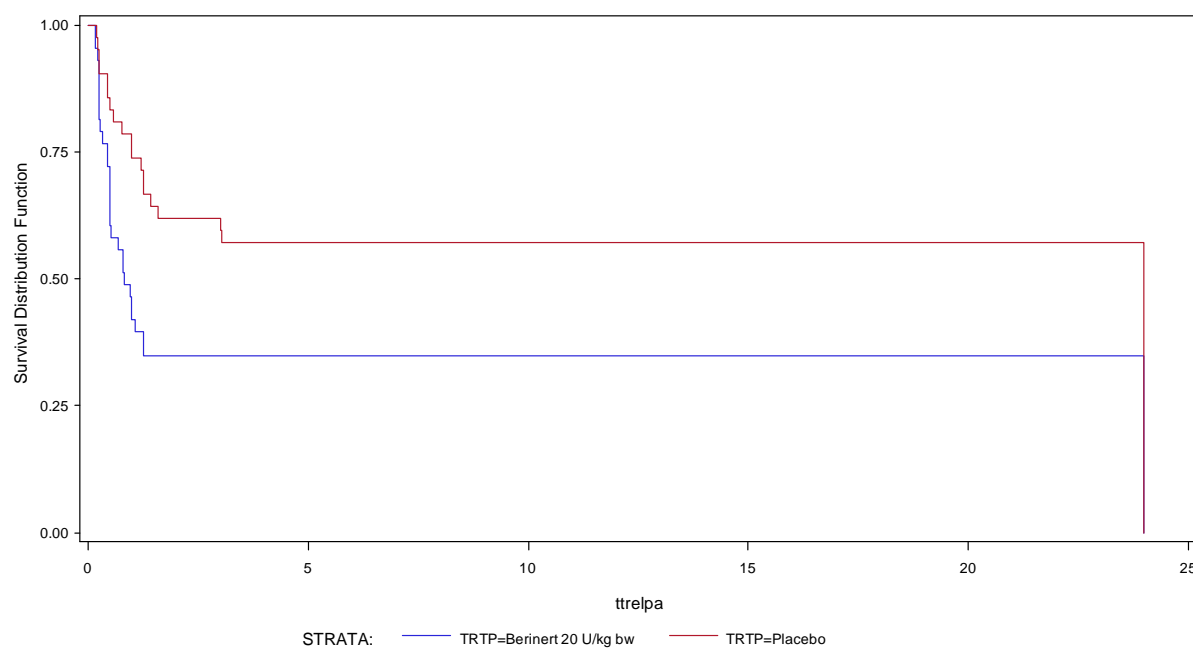
At FDA clinical reviewer's suggestion, this reviewer conducted three primary efficacy analyses and all three had a one-sided p-value less than 0.0249 (0.0053, 0.0016 and 0.0004 respectively using the Wilcoxon rank sum test with normal approximation). The one-sided p-values for both the secondary endpoint analyses are less than 0.1 (0.000084 and 0.033 respectively using Fisher's exact test and Wilcoxon rank sum test). According to the study protocol, the one-sided p-value for the final analysis needs to be less than 0.0249 to win the primary efficacy endpoint. Also, based on agreement between FDA and the sponsor, at least one one-sided p-value for the final analysis of the two secondary endpoints needs to be less than 0.1 to win the trial. The study results meet both criteria.

NOTE: There were protocol violations with respect to the interim analysis plan. There were two interim analyses in the trial. The first interim analysis did not follow the SAP: much less subjects were enrolled than what the plan suggested for the sample size re-adjustment (about 40 per arm instead of 100). This procedure was not justified and it was not clear how much influence this deviation had on the type I error rate. The second interim analysis was proposed after the study was already initiated. Without prospective plan to control overall type I error taking all the interim analyses into account, it is hard to justify that type I error is still under control in this complicated study design.

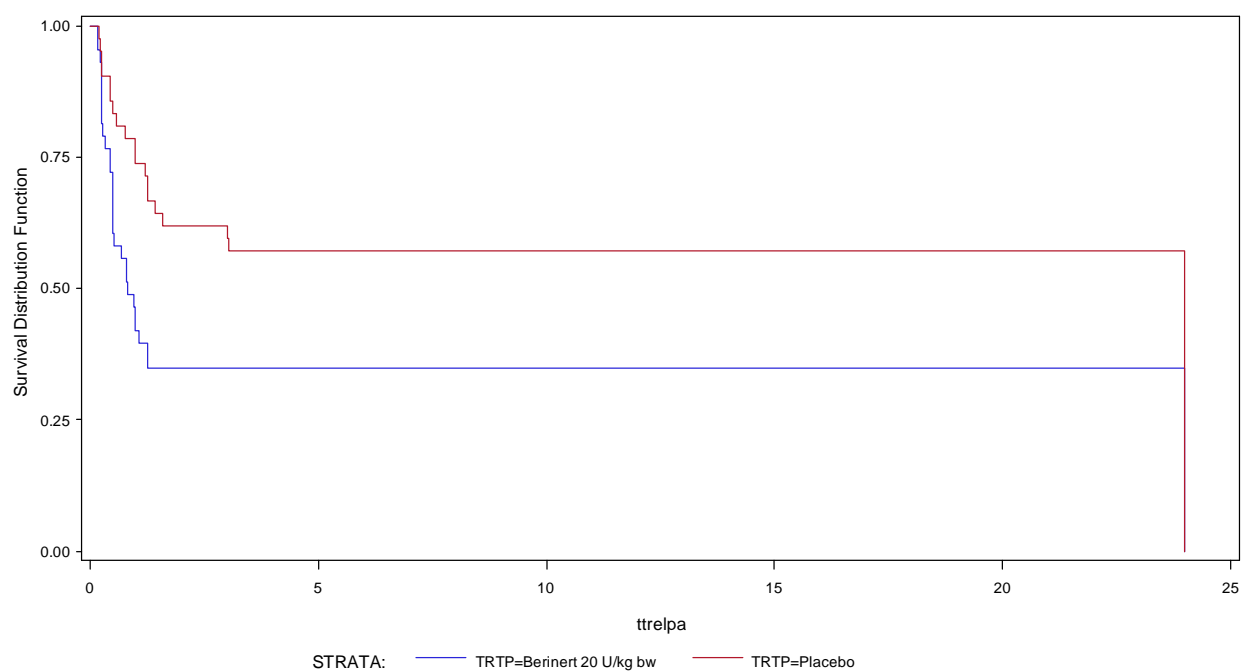
#### **APPENDIX**

The following are the Kaplan-Meier curves for the three analyses, described on page 8 of this review:

(a) Number of subjects rated poor/failure (24 hours): 15 (high dose Berinert) vs. 24 (placebo)



(b) Number of subjects rated poor/failure (24 hours): 13 (high dose Berinert) vs. 24 (placebo)



(R2) Number of subjects rated poor/failure (24 hours): 10 (high dose Berinert) vs. 23 (placebo)

