**DEPARTMENT OF HEALTH AND HUMAN SERVICES**

Public Health Service

---

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

# STATISTICAL REVIEW AND EVALUATION BLA

## SECOND CR FINAL MEMO

| | |
|---|---|
| **BLA/Supplement Number:** | 125426/0 |
| **Product Name:** | IB1001 |
| **Indication(s):** | Control and prevention of bleeding episodes and peri-operative management in patients with hemophilia B |
| **Final Applicant:** | Inspiration Biopharmaceuticals, INC. |
| **CBER Receipt Date**<br>**First CR Letter**<br>**CBER Receipt of First CR**<br>**Second CR Letter**<br>**CBER Receipt of Second CR** | 4/6/2012<br>2/1/2013<br>1/27/2014<br>7/29/2014<br>10/28/2014 |
| **Review Priority:** | Standard |
| **Statistical Branch:** | Therapeutics Evaluation Branch |
| **Primary Statistical Reviewer:** | Chunrong Cheng, Ph.D., Visiting Scientist |
| **Concurring Reviewer:** | Renee Rees, Ph.D., Team Leader |
| | Boguang Zhen, Ph.D., Branch Chief |
| **Committee Chair (CMC)** | Chava Kimchi-Sarfaty, Ph.D. |
| **Project Manager:** | Edward Thompson |

## Executive summary

This memo focuses on the review of a CMC issue raised in the second complete response (CR) letter. The applicant applied an inappropriate equivalence testing procedure to compare characteristics (b) (4) ) of the batches with (b)(4) , and batches with (b)(4) While this reviewer disagrees with the applicant's methodology, equivalence of all parameters can be concluded based on visual examination and summary statistics, without formal statistical hypothesis testing.

Regarding the one unresolved statistical comment, the applicant agreed to submit efficacy data on the original scale and has included the mean efficacy rates of the total annualized bleed rate (ABR) along with the median on the original scale in the package insert.

There are no unresolved statistical issues remaining in this submission.

## Background

The statistical review for the original BLA was completed on January 4, 2013, and a second statistical review (for the applicant's response to the first CR letter) was completed on July 14, 2014. That statistical review contains one statistical comment which was conveyed to the applicant in the second CR letter (dated July 29, 2014) as Item #12. It requested the applicant to submit the data on mean annualized bleeding rates (ABR) and any other efficacy measures using the original scale. This current (third) statistical review evaluates the applicant's response (dated October 28, 2014) to the second CR letter.

The applicant's response to Item #12 is satisfactory. The package insert has been revised to include the mean efficacy rates of the total ABR along with the median on the original scale.

In addition, Item 1b of the second CR letter requested the applicant to provide a complete characterization of batches with (b)(4) . In the response received on October 28, 2014, the applicant provided a report in Appendix 1 which assessed equivalency of batches with (b)(4) and batches with (b)(4) for (b)(4) drug product (DP). The CMC reviewer consulted this reviewer regarding the validity of the statistical tools and the calculations used in Appendix 1. Based on this reviewer's input, the CMC reviewer sent the following information request (IR) to the applicant on December 10, 2014:

1. With regard to your response to Item #1 of the CR letter:
    a. In your response to CR Item #1 you have used a value which you termed equivalence acceptance criteria (EAC). Please explain the rationale to determine the exact value and provide validation.

The response to the IR letter was received on December 22, 2014. Upon the request of the CMC reviewer, this reviewer conducted a formal review of the response, focusing on the above Item 1a.

*Note: to clarify, #1b refers to the item number in the second CR letter, while #1a refers to the item number in the IR made thereafter.*

## Summary of response to Item 1a

The applicant used a two one-sided t-test (TOST) to demonstrate statistical equivalency between two datasets. Defining $\overline{y_1}$ as the sample mean of dataset 1 (containing batches [redacted] and $\overline{y_2}$ as the sample mean of dataset 2 (containing batches with [redacted], then average equivalence is demonstrated if both of the following equations are satisfied:

$$\text{(i) } \overline{y_1} - \overline{y_2} + \text{ME} < \text{EAC}$$

$$\text{(ii) } \overline{y_1} - \overline{y_2} - \text{ME} > -\text{EAC},$$

where ME is the margin of error and is primarily determined by the standard deviation in dataset 1.

EAC is defined as below:

$$\text{(v) UEDL} = \overline{y_1} - \text{ME} + \text{EAC}$$

$$\text{(vi) LEDL} = \overline{y_1} + \text{ME} - \text{EAC}$$

*Note: for ease of reference the equation numbers in this review are from the applicant's response.*

UEDL and LEDL stand for the upper and lower equivalence determination limit respectively. The applicant uses the maximum (max) and the minimum (min) values in dataset 1 for UEDL and LEDL values respectively. The above approach has been described previously by Bower (1). However, Bower's article does not explain how EAC is derived but simply states that the utility of the chosen EAC can be objectively assessed before data from the new process are collected.

**Reviewer's derivation and re-parameterization of the equations**

1) Rewrite equations (v and vi) as

$$\text{(vii) EAC} = \text{UEDL} - \overline{y_1} + \text{ME}$$

$$\text{(viii) EAC} = \overline{y_1} + \text{ME} - \text{LEDL}$$

2) Replace the EAC in equation (ii) with equation (vii), we get

$$\overline{y_1} - \overline{y_2} - \text{ME} > -(\text{UEDL} - \overline{y_1} + \text{ME})$$

Since the ME and $\overline{y_1}$ are cancelled out, it becomes

$$\text{(ix) } \overline{y_2} < \text{UEDL}$$

3) Similarly, replace the EAC in equation (i) with equation (viii), we get

$$\overline{y_1} - \overline{y_2} + \text{ME} < (\overline{y_1} + \text{ME} - \text{LEDL})$$

, and it becomes

$$\text{(x) } \overline{y_2} > \text{LEDL}$$

4) Lastly, replace LEDL and UEDL in equations (ix) and (x) with "min" and "max" of dataset 1 respectively, we get

$$\text{(xi) } \min < \overline{y_2} < \max$$

After the above derivation and re-parameterization, it is clear that essentially the proposed testing procedure for equivalence is simply to access whether the mean of dataset 2 falls between the minimum and maximum values of dataset 1. The parameters ME and EAC do not play any real role indeed. Therefore, the following IR was sent to the sponsor to request revision of the test procedure.

**IR sent on January 22, 2015**

Regarding your response to Item 1a in the IR letter dated December 10, 2014, we agree that the two one-sided t-test (TOST) is an appropriate test to evaluate statistical equivalency between two datasets. However, if you replace the EAC value from equations v and vi (on page 2 of 12) into equations ii and i (on page 2 of 12), respectively, your testing procedure is simply reduced to assess whether the mean of dataset 2 falls between the LEDL (minimum of dataset 1) and UEDL (maximum of dataset 1). The parameters ME and EAC play none to minimal role indeed. This can be further confirmed from the example you provide on page 3 of 12.

We don't agree with this approach for the following reasons.

1. The range of dataset 1 may be very large in the presence of extreme values.
2. Comparing the mean of dataset 2 to the minimum and maximum values of dataset 1 is an extremely loose test. For example, in a rare situation, if 50% of dataset 2 points have values less than the minimum value of dataset 1, and the other 50% of dataset 2 points have values greater than the maximum value of dataset 1, then it is likely that equivalence will still be established.
3. It indeed does not take the variability of dataset 1 into account as ME gets cancelled out in the substitution mentioned above.

Therefore, your procedure poses a great risk to falsely claim equivalence for two potentially very different datasets. Please propose a more strict testing procedure which takes the data variability into account.

**Applicant's response to above IR**

The applicant's response to above IR was received on February 5, 2015. The response is summarized below.

- Based on visual examination and summary statistics, the applicant claims that that all parameters are highly similar for the test dataset (dataset 2) and the comparator dataset (dataset 1).
- The applicant maintains that the ME and the EAC are not cancelled out and are used in the calculation of statistical equivalency.

*Reviewer's evaluation: this reviewer agrees that equivalence can be concluded based on visual examination and summary statistics; however, the applicant did not provide new evidence to justify the equivalence testing procedure used.*

**Conclusion and recommendation**

1. Essentially, the proposed equivalence testing procedure only assesses whether the mean of dataset 2 (batches with (b)(4) falls between the minimum and maximum of dataset 1(batches with (b)(4) This approach is not acceptable.
2. FDA has conveyed to the applicant why this method is not acceptable, however, the applicant maintains that it is correct.
3. Since equivalence of the two datasets for all parameters can be concluded based on visual examination and summary statistics, formal statistical equivalence testing is not necessary in this case.
4. Caution should be exercised if the applicant applies the same equivalence testing procedure in future submissions.
5. There are no unresolved statistical issues remaining in this submission.

**Reference:**

Bower K M. Practical interpretation of equivalent acceptance criteria. Genetic Engineering and Biotechnology News. 2012 Feb 15:30.