CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

7610886Orig1s000

STATISTICAL REVIEW(S)

STATISTICAL REVIEW AND EVALUATION

Biometrics Division: VI

BLA NO.:	761088
SERIAL NO.:	0000
DATE RECEIVED BY THE CENTER:	May 29, 2018
DRUG NAME:	CT-P10 (proposed biosimilar to US-licensed Rituxan)
DOSAGE FORM:	Prefilled vials, 100 mg and 500 mg
INDICATIONS:	Non-Hodgkin's Lymphoma (NHL)
APPLICANT:	Celltrion Inc.
REVIEW FINISHED:	October 17, 2018
NAME OF PRIMARY STATISTICAL REVIEWER:	Yu-Ting Weng
NAME OF SECONDARY STATISTICAL REVIEWER:	Meiyu Shen
NAME OF PROJECT MANAGER:	Esther Park

_____

Reviewer: Yu-Ting Weng, Ph.D., Mathematical Statistician, CDER/OTS/OB/DB VI
Secondary reviewer: Meiyu Shen, Ph.D., Team Leader, CDER/OTS/OB/DB VI

Concur: _____
Yi Tsong, Ph.D., Division Director, DBVI

Distribution: CDER/OTS/OB/DB VI/Yi Tsong
CDER/OTS/OB/DB VI/Meiyu Shen
CDER/OTS/OB/Lillian Patrician
CDER/OBP/Nina Brahme
CDER/OBP/Haoheng Yan
CDER/OBP/Bazarragchaa Damdinsuren
CDER/OMPT/CDER/OND/OHOP/DHP/Esther Park
CDER/TBBS

TABLE OF CONCENTRATIONS

List of Tables

List of Figures

1 INTRODUCTION

Due to the manufacturing facility inspection issue, the FDA issued a warning letter on January 26, 2018 and a complete response letter on February 28, 2018. On May 29, 2018, Celltrion (the applicant) resubmitted to the US Food and Drug Administration (FDA) a 351(k) BLA which included analytical similarity assessments of comparing CT-P10, US-licensed Rituxan (US), and EU-approved Rituxan (EU).

Two Tier 1 Quality Attributes (QAs) are relative activities of CDC and ADCC and the analytical similarity data is the same as the one submitted on April 28, 2017.

The FDA CMC statistics reviewers carefully evaluated data for the relative activities of CDC and ADCC provided in the BLA resubmission. The applicant's statistical equivalence testing (Tier 1 approach) is provided in Section 3, and CMC statistical reviewer's independent statistical equivalence testing analyses using modified Wald test are present in Section 4.

2 APPLICANT'S STATISTICAL EQUIVALENCE TESTING

In this resubmission, the applicant followed the FDA's recommendation to conduct Tier 1 statistical equivalence testing with the margin defined as $1.5\hat{\sigma}_R$, where $\hat{\sigma}_R$ is the sample standard deviation based on the reference product lots, for the relative CDC activity and the relative ADCC activity at three concentration levels.

The CMC statistical reviewer performs an independent statistical analysis in the next section.

3 FDA STATISTICAL ANALYSES

To evaluate analytical similarity, the FDA recommended the applicant to apply a tiered approach in the FDA responses to IND meetings with the applicant. That is, product QAs amenable to statistical evaluation are assigned to three tiers based on their criticality. The QAs with potential highest risk in product quality, efficiency, safety, and PK/PD are generally assigned to Tier 1, in which analytical similarity is assessed by statistical equivalence test. QAs with lower impact are generally assigned to Tier 2 and their analytical similarity is evaluated by Quality Range approach. That is, a high percentage of the biosimilar data should be covered by $(\hat{\mu}_R - X\hat{\sigma}_R, \hat{\mu}_R + X\hat{\sigma}_R)$, where $\hat{\mu}_R$ is the sample mean, $\hat{\sigma}_R$ is the sample standard deviation based on the reference product lots, and the multiplier X typically ranges from 2 to 4. The QAs with the lowest risk are generally assigned to Tier 3 and their analytical similarity is evaluated by side-by-side comparison using graphic display.

This review focuses on the equivalence test in Tier 1.

3.1 Data analyzed

The applicant resubmitted the analytical data on May 29, 2018 and the data is the same as the one submitted on April 28, 2017. The summary of CT-P10 lots for Tier 1 analysis is provided in Table 1. The CMC statistical reviewer conducted Tier 1 statistical equivalence testing based on these 15 selected independent lots for each QA.

Table 1. Summary of CT-P10 Lots for Tier 1 Analysis

DP lot number	DS lot number	Tested lots for Tier 1 QAs	
		CDC	ADCC
	E13220C001		
13C1C004	13220C002	X	X
13C1C003	13220C003	X	X
14C1C001	13220C004	X	X
	13220C005		
	13220C006		
	13220C007		
	13220C008		
	13220C009		
	13220C011		
14C2C002	14230C001	X	X
14C2C003	14230C002	X	X
14C2C004	14230C003	X	X
15C2C01	15230C001	X	X
15C2C02	15230C002	X	X
15C2C08	15230C003	X	X
15C2C09	15230C004	X	X
15C2C10	15230C005	X	X
	15230C006		
15C2C03	15230C007	X	X
15C2C04	15230C009	X	X
15C2C05	15230C010	X	X
15C2C06	15230C011	X	X

X: tested independent lots included in the analysis

3.2 Statistical method

Let μ_T and μ_R be the population mean of the QA for the test product and the population mean of the QA for the reference product, respectively. Let σ_R be the standard deviation of the QA of interest for the reference product. To conclude the equivalence in the QA of interest between the test product and the reference product, we aim to reject the null hypothesis of the following null and alternative hypotheses (Tsong et al., 2017) [1]:

$$H_0 : \mu_T - \mu_R \le \theta_1 \text{ or } \mu_T - \mu_R \ge \theta_2$$
$$H_1 : \theta_1 < \mu_T - \mu_R < \theta_2$$

where $\theta_1 = -1.5\sigma_R$, $\theta_2 = 1.5\sigma_R$, θ_1 and θ_2 are equivalence margins.

In the current practice for fixed margin, we reject H_0 if 90% CI for the mean difference in the QA of interest falls within $(-1.5\sigma_R, 1.5\sigma_R)$. In other words, we conclude that the equivalence in the QA of interest between the test product and the reference product if 90% CI for the mean difference in the QA of interest falls within $(-1.5\sigma_R, 1.5\sigma_R)$. This specific equivalence margin was set as 1.5 times the standard deviation of the QA for the reference product to ensure an adequate power for the case in which a small but sufficient number of lots are available for testing. For example, the probability of rejecting H_0 in the above two one-sided tests procedure with the equivalence margin being $\pm 1.5\sigma_R$ is 87% if the true mean difference is $0.125\sigma_R$ for a sample size of 10 test product lots and 10 reference product lots.

First, we estimate σ_R by the sample variability of the reference product, and then θ_1 and θ_2 are treated as a constant, but not a random variable in the statistical analysis.

Let X_{Tj} be the observed value of the QA of interest for Lot *j* of the test product (the proposed biosimilar product) and X_{Rj} be the observed value of the QA of interest for Lot *j* of the reference product. Since the two products are manufactured by two manufacturers, two products are independent. $\bar{X}_i = \sum_{j=1}^{n_i} X_{ij} \Big/ n_i$, and $S_i^2 = \sum_{j=1}^{n_i} \left(X_{ij} - \bar{X}_i\right) \Big/ \left(n_i - 1\right)$, where n_i is the number of lots in the i^{th} product, $i = T, R$.

Under the unequal variance of the test product and the reference product, the (1-2α)*100% CI of the mean difference in the QA of interest can be calculated as:

$$\left(\bar{X}_T - \bar{X}_R - t_\alpha(\nu)\sqrt{\frac{S_T^2}{n_T} + \frac{S_R^2}{n_R}},\ \bar{X}_T - \bar{X}_R + t_\alpha(\nu)\sqrt{\frac{S_T^2}{n_T} + \frac{S_R^2}{n_R}} \right).$$

where $t_\alpha(\nu)$ is the 1-α quantile and ν is the degrees of freedom calculated by Satterthwaite's approximation.

If $n_R > 1.5n_T$, the (1-2α)*100% sample size imbalanced adjusted CI of the mean difference in the QA of interest can be calculated as (Dong et al., 2017a) [2]:

$$\left(\bar{X}_T - \bar{X}_R - t_\alpha(\nu^*)\sqrt{\frac{S_T^2}{n_T} + \frac{S_R^2}{n_R^*}},\ \bar{X}_T - \bar{X}_R + t_\alpha(\nu^*)\sqrt{\frac{S_T^2}{n_T} + \frac{S_R^2}{n_R^*}} \right).$$

where $n_R^* = \min(n_R, 1.5n_T)$ and $\nu^* = \dfrac{\left(\frac{S_T^2}{n_T} + \frac{S_R^2}{n_R^*}\right)^2}{\frac{1}{n_T - 1}\left(\frac{S_T^2}{n_T}\right)^2 + \frac{1}{n_R - 1}\left(\frac{S_R^2}{n_R^*}\right)^2}$.

If $n_T > 1.5n_R$, we can apply a similar approach as above with $n_T^* = \min(1.5 \times n_R, n_T)$ for the CI calculation. In the following analyses, we use α=0.05.

However, the current practice for fixed margin results in inflating type I error rate and reducing power when sample size is not sufficiently large, as pointed out by Dong et al. (2017a) [2]. Hence, one needs to consider σ_R as a parameter to be estimated with the study data. We recommend the MWCMLE method for addressing the issue.

The above null and alternative hypotheses for reference scaled equivalence can be rewritten as the two-one-sided hypothesis testing as follows.

$$H_{0U}: \mu_T - \mu_R \geq f\sigma_R \text{ vs. } H_{aU}: \mu_T - \mu_R < f\sigma_R$$

$$H_{0L}: \mu_T - \mu_R \leq -f\sigma_R \text{ vs. } H_{aL}: \mu_T - \mu_R > -f\sigma_R$$

where μ_T and μ_R are the population means of the test product and the reference product, respectively; σ_R is the Standard Deviation (SD) of the reference product; and f is 1.7.

The test statistics of the MWCMLE method to test H_{0U} and H_{0L} in (1) are as follows:

$$W_L = \frac{\hat{\mu}_T - \hat{\mu}_R + f\tilde{\sigma}_R}{\sqrt{\frac{\breve{\sigma}^2_{T_L}}{n_T} + \left(\frac{1}{n_R} + \frac{f^2 V_{n_R}}{n_R - 1}\right)\breve{\sigma}^2_{R_L}}} \text{ and } W_U = \frac{\hat{\mu}_T - \hat{\mu}_R - f\tilde{\sigma}_R}{\sqrt{\frac{\breve{\sigma}^2_{T_U}}{n_T} + \left(\frac{1}{n_R} + \frac{f^2 V_{n_R}}{n_R - 1}\right)\breve{\sigma}^2_{R_U}}},$$

where $\hat{\mu}_T$, $\hat{\mu}_R$, and $\tilde{\sigma}_R$ denote the MLE for the mean of the test product, the MLE for the mean of the reference product, and the sample estimator for the SD of the reference product, respectively. When X_T follows $N(\mu_T, \sigma_T^2)$ distribution and X_R follows $N(\mu_R, \sigma_R^2)$ distribution, $\hat{\mu}_T = \frac{\sum_{i=1}^{n_T} X_{T_i}}{n_T} = \bar{X}_T$, $\hat{\mu}_R = \frac{\sum_{i=1}^{n_R} X_{R_i}}{n_R} = \bar{X}_R$, $\tilde{\sigma}_R = \sqrt{\frac{\sum_{i=1}^{n_R}\left(X_{R_i} - \hat{\mu}_R\right)^2}{n_R - 1}}$ where n_T and n_R denote the sample sizes of test product lots and reference lots, respectively. σ_T^2 is the variance of the test product. Under the null hypothesis, each test follows an asymptotically standard normal distribution. The overall null hypothesis is rejected by either of the following two criteria: first, $W_L > Z_{1-\alpha}$ and $W_U < -Z_{1-\alpha}$ at significance level α, where Z_p is the 100pth percentile of the standard normal distribution; or second, the intersection (L, U) of two one-sided (1- α)*100% confidence intervals falls within the equivalence margin $(-f\tilde{\sigma}_R, f\tilde{\sigma}_R)$ and is listed as follows.

$$(L, U) =$$

$$\left(\hat{\mu}_T - \hat{\mu}_R - Z_{1-\alpha}\sqrt{\frac{\breve{\sigma}^2_{T_L}}{n_T} + \left(\frac{1}{n_R} + \frac{f^2 V_{n_R}}{n_R - 1}\right)\breve{\sigma}^2_{R_L}}, \hat{\mu}_T - \hat{\mu}_R + Z_{1-\alpha}\sqrt{\frac{\breve{\sigma}^2_{T_U}}{n_T} + \left(\frac{1}{n_R} + \frac{f^2 V_{n_R}}{n_R - 1}\right)\breve{\sigma}^2_{R_U}}\right)$$

Following the above same logic of the sample size imbalanced adjustment, the Adjusted MWCMLE statistics for imbalanced sample (AMWCMLE) can be written as follows:

$$W_{L'} = \frac{\hat{\mu}_T - \hat{\mu}_R + f\tilde{\sigma}_R}{\sqrt{\frac{\breve{\sigma}^2_{T_L}}{n_T} + \left(\frac{1}{n_R^*} + \frac{f^2 V_{n_R}}{n_R - 1}\right)\breve{\sigma}^2_{R_L}}} \text{ and } W_{U'} = \frac{\hat{\mu}_T - \hat{\mu}_R - f\tilde{\sigma}_R}{\sqrt{\frac{\breve{\sigma}^2_{T_U}}{n_T} + \left(\frac{1}{n_R^*} + \frac{f^2 V_{n_R}}{n_R - 1}\right)\breve{\sigma}^2_{R_U}}},$$

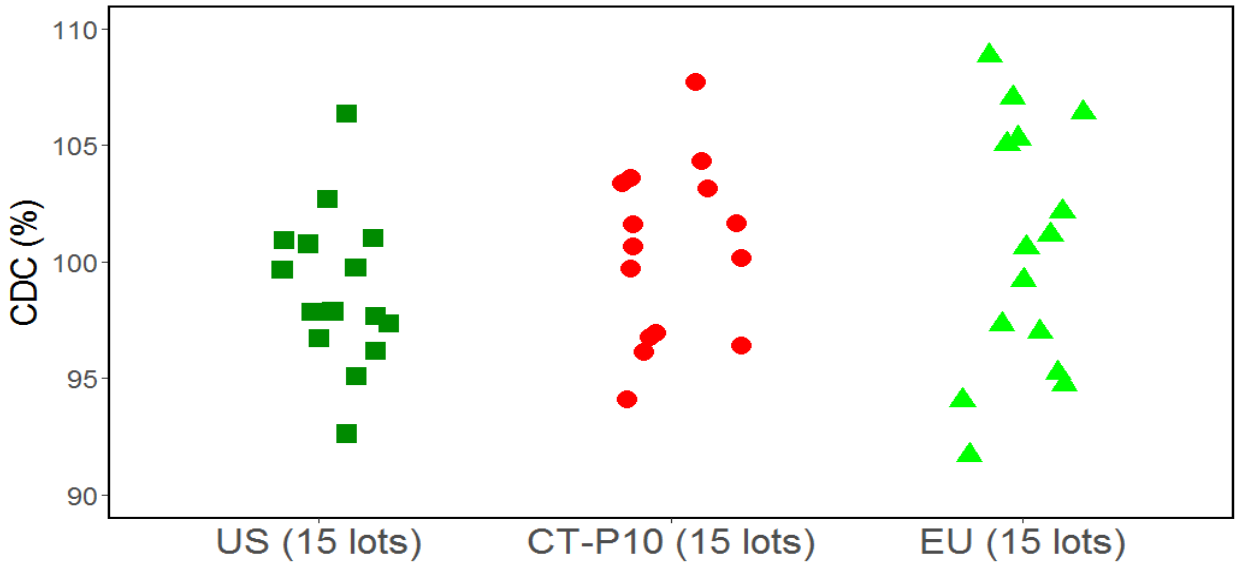
where the sample sizes of adjusted reference lots, n_R^* , is equal to $\min(1.5 \times n_T, n_R)$. n_T and n_R denote the sample sizes of test product lots and the sample sizes of reference product lots, respectively.

More details are described in Weng et al. (2018) [3].

3.3 FDA statistical equivalence testing for the relative CDC activity

The relative CDC activity's data points of CT-P10, US, and EU are displayed in Figure 1. CT-P10 has the largest sample mean among three products. The sample variability of CT-P10 is larger than the sample variability of US; the sample variability of CT-P10 is smaller than the sample variability of EU.

Figure 1. Scatter plot of the relative CDC activity for CT-P10, US-licensed Rituxan, and EU-approved Rituxan



Fifteen lots of all three products were included for the statistical equivalence testing for the relative CDC activity. Descriptive statistics for the relative CDC activity's data are listed in Table 2.

Table 2. Descriptive statistics for the relative CDC activity data

Product	Number of lots	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
CT-P10	15	100.47	3.83	94	108
US	15	98.93	3.28	93	106
EU	15	100.33	5.27	92	109

From Table 3, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative CDC activity, using the current practice for fixed margin, pass equivalence testing.

Table 3. Equivalence testing results for the relative CDC activity using the current practice for fixed margin

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	1.53	(-0.69, 3.75)	(-4.93, 4.93)	Yes
CT-P10 vs. EU	(15, 15)	0.13	(-2.74, 3.01)	(-7.91, 7.91)	Yes
EU vs. US	(15, 15)	1.40	(-1.35, 4.15)	(-4.93, 4.93)	Yes

From Table 4, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative CDC activity, using the MWCMLE method for parameter margin, also pass equivalence testing.

Table 4. Equivalence testing results for the relative CDC activity using the MWCMLE for parameter margin

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	1.53	(-1.68, 4.08)	(-5.58, 5.58)	Yes
CT-P10 vs. EU	(15, 15)	0.13	(-3.61, 3.81)	(-8.96, 8.96)	Yes
EU vs. US	(15, 15)	1.40	(-2.21, 4.47)	(-5.58, 5.58)	Yes

3.4 FDA statistical equivalence testing for the relative ADCC activity

The applicant submitted the relative ADCC activity data at three concentration levels: 0.01 μg/mL, 0.035 μg/mL, and 0.122 μg/mL. We evaluate the data at each concentration level as follows.

3.4.1 The relative ADCC activity at 0.01 μg/mL

The relative ADCC activity's data points of CT-P10, US, and EU at 0.01 μg/mL are displayed in Figure 2. CT-P10 has the largest sample mean and sample variability among three products.

Fifteen lots of all three products were included for the statistical equivalence testing for the relative ADCC activity at 0.01 μg/mL. Descriptive statistics for the relative ADCC activity's data at 0.01 μg/mL are listed in Table 5.

From Table 6, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative ADCC activity at 0.01 μg/mL, using the current practice for fixed margin, pass equivalence testing.

Figure 2. Scatter plot of the relative ADCC activity for CT-P10, US-licensed Rituxan, and EU-approved Rituxan at 0.01 µg/mL

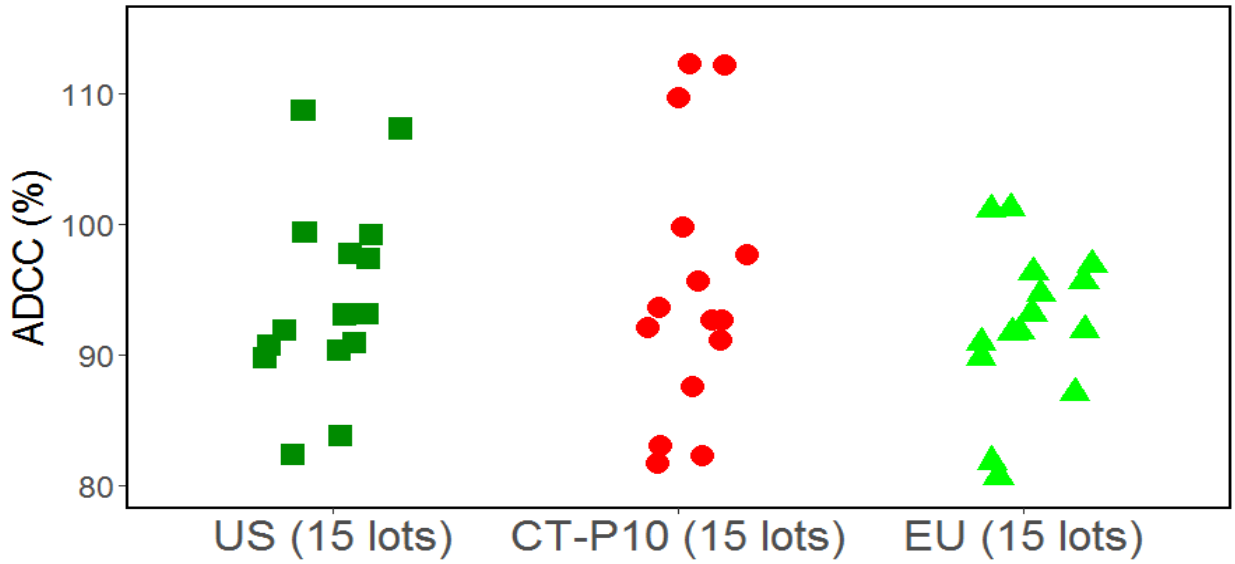


Table 5. Descriptive statistics for the relative ADCC activity data at 0.01 µg/mL

Product	Number of lots	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
CT-P10	15	95.07	10.01	82	112
US	15	94.33	7.40	82	109
EU	15	92.4	5.84	81	101

Table 6. Equivalence testing results for the relative ADCC activity at 0.01 µg/mL using the current practice for fixed margin

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	0.73	(-4.75, 6.22)	(-11.11, 11.11)	Yes
CT-P10 vs. EU	(15, 15)	2.67	(-2.47, 7.80)	(-8.76, 8.76)	Yes
EU vs. US	(15, 15)	-1.93	(-6.08, 2.22)	(-11.11, 11.11)	Yes

From Table 7, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative ADCC activity at 0.01 µg/mL, using the MWCMLE method for parameter margin, also pass equivalence testing.

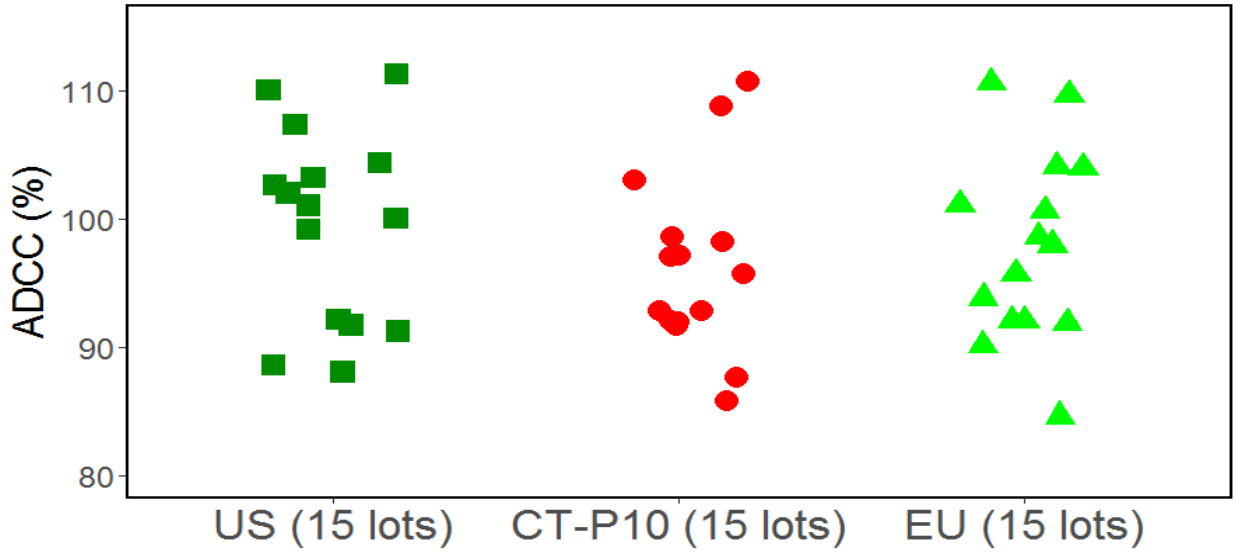
Table 7. Equivalence testing results for the relative ADCC activity at 0.01 µg/mL using the MWCMLE for parameter margin

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	0.73	(-6.19, 7.35)	(-12.59, 12.59)	Yes
CT-P10 vs. EU	(15, 15)	2.67	(-4.00, 8.33)	(-9.93, 9.93)	Yes
EU vs. US	(15, 15)	-1.93	(-6.89, 4.01)	(-12.59, 12.59)	Yes

3.4.2 The relative ADCC activity at 0.035 µg/mL

The relative ADCC activity's data points of CT-P10, US, and EU at 0.035 µg/mL are displayed in Figure 3. CT-P10 has the smallest sample mean and sample variability among three products.

Figure 3. Scatter plot of the relative potency of the ADCC assay for CT-P10, US-licensed Rituxan, and EU-approved Rituxan at 0.035 µg/mL



Fifteen lots of all three products were included for the statistical equivalence testing for the relative ADCC activity at 0.035 µg/mL. Descriptive statistics for the relative ADCC activity's data at 0.035 µg/mL are listed in Table 8.

From Table 9, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative potency of the ADCC assay at 0.035 µg/mL, using the current practice for fixed margin, pass equivalence testing.

From Table 10, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative potency of the ADCC assay at 0.035 µg/mL, using the MWCMLE method for parameter margin, also pass equivalence testing.

Table 8. Descriptive statistics for the relative ADCC activity data at 0.035 µg/mL

Product	Number of lots	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
CT-P10	15	96.4	7.00	86	111
US	15	99.47	7.46	88	111
EU	15	97.93	7.39	85	111

Table 9. Equivalence testing results for the relative ADCC activity at 0.035 µg/mL using the current practice for fixed margin

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	-3.07	(-7.56, 1.43)	(-11.19, 11.19)	Yes
CT-P10 vs. EU	(15, 15)	-1.53	(-6.00, 2.94)	(-11.09, 11.09)	Yes
EU vs. US	(15, 15)	-1.53	(-6.15, 3.08)	(-11.19, 11.19)	Yes

Table 10. Equivalence testing results for the relative ADCC activity at 0.035 µg/mL using the MWCMLE for parameter margin

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	-3.07	(-8.31, 3.62)	(-12.69, 12.69)	Yes
CT-P10 vs. EU	(15, 15)	-1.53	(-6.96, 4.63)	(-12.57, 12.57)	Yes
EU vs. US	(15, 15)	-1.53	(-7.13, 4.78)	(-12.69, 12.69)	Yes

3.4.3 The relative ADCC activity at 0.122 µg/mL

The relative ADCC activity's data points of CT-P10, US, and EU at 0.122 µg/mL are displayed in Figure 4. CT-P10 has the smallest sample mean among three products. In addition, CT-P10 has the largest sample variability among three products.

Fifteen lots of all three products were included for the statistical equivalence testing for the relative ADCC activity at 0.122 µg/mL. Descriptive statistics for the relative ADCC activity's data at 0.122 µg/mL are listed in Table 11.

From Table 12, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative ADCC activity at 0.122 µg/mL, using the current practice for fixed margin, pass equivalence testing.

Figure 4. Scatter plot of the relative ADCC activity for CT-P10, US-licensed Rituxan, and EU-approved Rituxan at 0.122 µg/mL

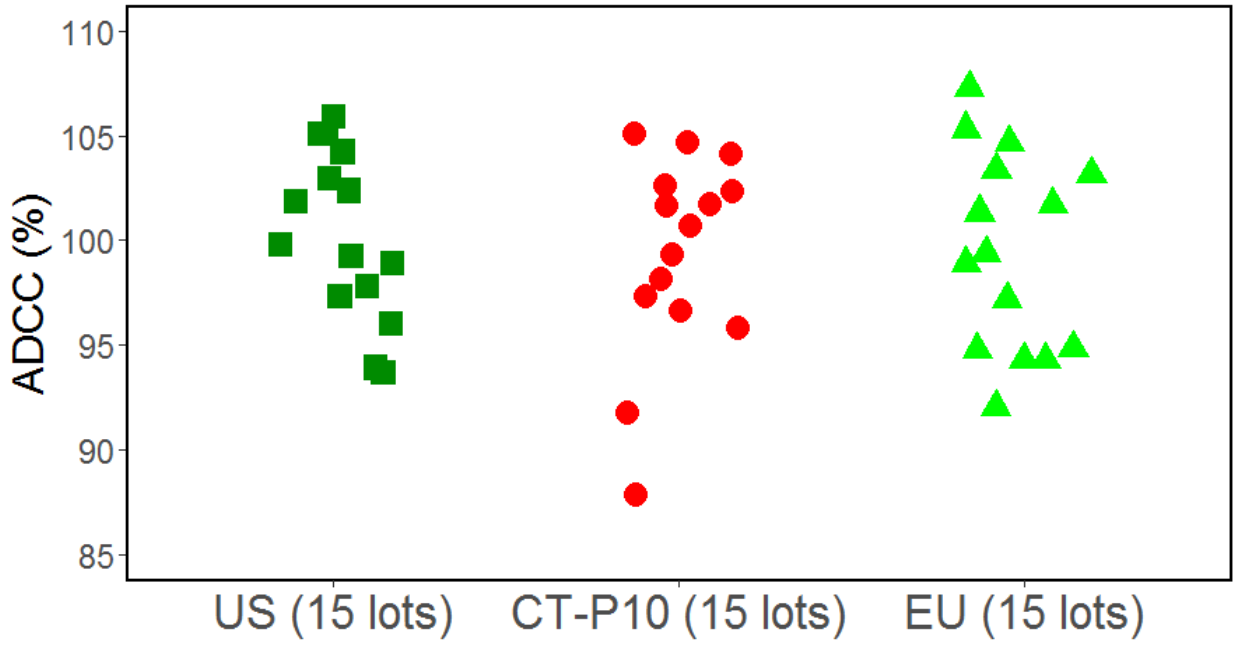


Table 11. Descriptive statistics for the relative ADCC activity data at 0.122 µg/mL

Product	Number of lots	Sample mean, mg/mL	Sample standard deviation, mg/mL	Minimum, mg/mL	Maximum, mg/mL
CT-P10	15	99.4	4.85	88	105
US	15	100.27	3.97	94	106
EU	15	99.4	4.73	92	107

Table 12. Equivalence testing results for the relative ADCC activity at 0.122 µg/mL using the current practice for fixed margin

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	-0.87	(-3.62, 1.89)	(-5.96, 5.96)	Yes
CT-P10 vs. EU	(15, 15)	0.00	(-2.98, 2.98)	(-7.10, 7.10)	Yes
EU vs. US	(15, 15)	-0.87	(-3.58, 1.85)	(-5.96, 5.96)	Yes

From Table 13, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative ADCC activity at 0.122 μg/mL, using the MWCMLE method for parameter margin, also pass equivalence testing.

Table 13. Equivalence testing results for the relative ADCC activity at 0.122 μg/mL using the MWCMLE for parameter margin

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	-0.87	(-4.15, 2.80)	(-6.75, 6.75)	Yes
CT-P10 vs. EU	(15, 15)	0.00	(-3.80, 3.80)	(-8.05, 8.05)	Yes
EU vs. US	(15, 15)	-0.87	(-4.11, 2.76)	(-6.75, 6.75)	Yes

4 CONCLUSION AND RECOMMENDATION

The results from the statistical equivalence analyses for the relative activities of CDC and ADCC at three concentration levels supported a demonstration that the proposed biosimilar CT-P10 is highly similar to US-licensed Rituxan.

5 REFERENCE

[1] Tsong Y., Dong X., Shen M. (2017): Development of statistical methods for analytical similarity assessment, Journal of Biopharmaceutical Statistics: 27:2, pages 308-316

[2] Dong X., Weng Y.T., Tsong Y. (2017b). Adjustment for unbalanced sample size for analytical biosimilar equivalence assessment. Journal of Biopharmaceutical Statistics: 27:2, pages 220-232

[3] Weng Y.T., Tsong Y., Shen M., Wang C. (2018). Improved Wald Test for Reference Scaled Equivalence Assessment of Analytical Biosimilarity. Accepted by International Journal of Clinical Biostatistics and Biometrics. DOI: 10.23937/2469-5831/1510016

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

YU-TING WENG
10/18/2018

MEIYU SHEN
10/18/2018

YI TSONG
10/19/2018

STATISTICAL REVIEW AND EVALUATION

Biometrics Division: VI

BLA NO.:	761088
SERIAL NO.:	0000
DATE RECEIVED BY THE CENTER:	April 28, 2017
DRUG NAME:	CT-P10 (proposed biosimilar to US-licensed Rituxan)
DOSAGE FORM:	Prefilled vials, 100 mg and 500 mg
INDICATIONS:	(b) (4)
APPLICANT:	Celltrion Inc.
REVIEW FINISHED:	Jan 5, 2018
NAME OF PRIMARY STATISTICAL REVIEWER:	Yu-Ting Weng
NAME OF SECONDARY STATISTICAL REVIEWER:	Meiyu Shen
NAME OF PROJECT MANAGER:	Esther Park

_____

Reviewer: Yu-Ting Weng, Ph.D., Mathematical Statistician, CDER/OTS/OB/DB VI
Secondary reviewer: Meiyu Shen, Ph.D., Team Leader, CDER/OTS/OB/DB VI

Concur: _____
Yi Tsong, Ph.D., Division Director, DBVI

Distribution: CDER/OTS/OB/DB VI/Yi Tsong
CDER/OTS/OB/DB VI/Meiyu Shen
CDER/OTS/OB/Lillian Patrician
CDER/OBP/Nina Brahme
CDER/OBP/Bazarragchaa Damdinsuren
CDER/OMPT/CDER/OND/OHOP/DHP/Esther Park
CDER/TBBS

TABLE OF CONCENTRATIONS

List of Tables

List of Figures

1 PRIMARY REVIEWER SUMMARY RECOMMENDATION (CMC STATISTICS)

The following lists the summary of recommendations from Dr. Nina Brahme, OBP CMC reviewer.

Due to facility compliance issues, a pre-license inspection could not be performed. Pending an inspection that includes an assessment of similarity data integrity, the data provided support the conclusion that CT-P10 is highly similar to US-licensed Rituxan.

CMC statistical reviewer's recommendation

The results of the statistical equivalence analyses for the Tier 1 Quality Attributes (QAs) support a demonstration that CT-P10 is highly similar to US-licensed Rituxan. However, the validity of the conclusions of the statistical equivalence analyses for Tier 1 QAs depends on the outcomes of the pending inspection which will assess the integrity of the analytical similarity data in analytical similarity studies.

2 INTRODUCTION

On April 28, 2017, Celltrion (the applicant) submitted to the US Food and Drug Administration (FDA) a 351(k) BLA which included analytical similarity assessments of comparing CT-P10, US-licensed Rituxan (US), and EU-approved Rituxan (EU).

Two Tier 1 QAs are relative potency of the CDC assay and the relative potency of the ADCC assay. The CMC statistical reviewer found out the following review issues and discussed internally with OBP CMC reviewers.

Issue 1. Section 3.2.S.5: the file titled "Reference Standards or Materials", for analytical results for CT-P10 primary reference standard (b) (4), it might not be a common practice that the applicant used a mixture of 5 lots of EU-approved Rituxan but not US-licensed Rituxan as the reference to define the biological activity.

Issue 2. Section 3.2.R: the file titled "Regional Information", for two Tier 1 QAs, the applicant defined the relative potency as the ratio of cell cytotoxicity but not the ratio of dose achieving a certain effect.

Issue 3. Section 3.2.R: the file titled "Regional Information", the applicant used the individual cytotoxicity at each concentration level for the relative potency of the ADCC assay. However, the applicant used the mean cytotoxicity at each concentration level for the relative potency of the CDC assay.

On September 21, 2017, OBP CMC team leader, Dr. Bazarragchaa Damdinsuren, emailed the following responses to the CMC statistical reviewer in the message titled "Re: BLA 761088: Defintion of relative potency of two Tier 1 QAs":

- Agree with the mixture of 5 lots of EU-approved Rituxan as the reference standard.
- Agree with the different definition of the relative potency.
- The reason that the use of the individual cytotoxicity at each concentration level for the relative potency of the ADCC with PBMC assay but not for others might due to internal

variability of the assay to target the linear slope of the response where the assay uses donor blood effector cells.

Due to the pending pre-license inspection issue, the FDA did not send an IR to the applicant for more clarification.

The FDA CMC statistics reviewers carefully evaluated data for the relative potency of the CDC assay and the relative potency of the ADCC assay provided in the initial BLA submission. The applicant's statistical equivalence testing (Tier 1 approach) is provided in Section 3, and CMC statistical reviewer's independent statistical equivalence testing analyses are present in Section 4.

3 APPLICANT'S STATISTICAL EQUIVALENCE TESTING

In this submission, the applicant followed the FDA's recommendation to conduct Tier 1 statistical equivalence testing with the margin defined as $1.5\hat{\sigma}_R$, where $\hat{\sigma}_R$ is the sample standard deviation based on the reference product lots, for the relative potency of the CDC assay and the relative potency of the ADCC assay at three concentration levels. Hence, if the analytical data were valid, the applicant's analyses would be valid. If the analytical data were not valid, then the applicant's analyses would not be valid.

The CMC statistical reviewer performs an independent statistical analysis in the next section.

4 FDA STATISTICAL ANALYSES

To evaluate analytical similarity, the FDA recommended the applicant to apply a tiered approach in the FDA responses to IND meetings with the applicant. That is, product QAs amenable to statistical evaluation are assigned to three tiers based on their criticality. The quality attributes with potential highest risk in product quality, efficiency, safety, and PK/PD are generally assigned to Tier 1, in which analytical similarity is assessed by statistical equivalence test. QAs with lower impact are generally assigned to Tier 2 and their analytical similarity is evaluated by Quality Range approach. That is, a high percentage of the biosimilar data should be covered by $(\hat{\mu}_R - X\hat{\sigma}_R, \hat{\mu}_R + X\hat{\sigma}_R)$, where $\hat{\mu}_R$ is the sample mean, $\hat{\sigma}_R$ is the sample standard deviation based on the reference product lots, and the multiplier X typically ranges from 2 to 4. The QAs with the lowest risk are generally assigned to Tier 3 and their analytical similarity is evaluated by side-by-side comparison using graphic display. More details are described in the FDA Draft Guidance on Analytical Similarity (2017) [1]

This review focuses on the equivalence test in Tier 1.

4.1 Data analyzed

The applicant submitted the analytical data on April 28, 2017. The summary of CT-P10 lots for Tier 1 analysis is provided in Table 1. Although the applicant selected a subset of all available lots for both Tier 1 QAs without providing a selection criterion, the OBP CMC reviewers agreed with the applicant's selection as indicated in the email on Sep 21, 2017 (for details, see Section 2). Thus, the CMC statistical reviewer conducted Tier 1 statistical equivalence testing based on these 15 selected independent lots for each QA.

Table 1. Summary of CT-P10 Lots for Tier 1 Analysis

DP lot number	DS lot number	Tested lots for Tier 1 QAs	
		CDC	ADCC
	E13220C001		
13C1C004	13220C002	X	X
13C1C003	13220C003	X	X
14C1C001	13220C004	X	X
	13220C005		
	13220C006		
	13220C007		
	13220C008		
	13220C009		
	13220C011		
14C2C002	14230C001	X	X
14C2C003	14230C002	X	X
14C2C004	14230C003	X	X
15C2C01	15230C001	X	X
15C2C02	15230C002	X	X
15C2C08	15230C003	X	X
15C2C09	15230C004	X	X
15C2C10	15230C005	X	X
	15230C006		
15C2C03	15230C007	X	X
15C2C04	15230C009	X	X
15C2C05	15230C010	X	X
15C2C06	15230C011	X	X

X: tested independent lots included in the analysis

4.2 Statistical method

Let μ_T and μ_R be the population mean of the QA for the test product and the population mean of the QA for the reference product, respectively. Let σ_R be the standard deviation of the QA of interest for the reference product. To conclude the equivalence in the QA of interest between the test product and the reference product, we aim to reject the null hypothesis of the following null and alternative hypotheses:

$$H_0 : \mu_T - \mu_R \le \theta_1 \text{ or } \mu_T - \mu_R \ge \theta_2$$
$$H_1 : \theta_1 < \mu_T - \mu_R < \theta_2$$

where $\theta_1 = -1.5\sigma_R$, $\theta_2 = 1.5\sigma_R$, θ_1 and θ_2 are equivalence margins.

We reject H_0 if 90% CI for the mean difference in the QA of interest falls within $(-1.5\sigma_R, 1.5\sigma_R)$. In other words, we conclude that the equivalence in the QA of interest between the test product and the reference product if 90% CI for the mean difference in the QA of interest falls within $(-1.5\sigma_R, 1.5\sigma_R)$. This specific equivalence margin was set as 1.5 times the standard deviation of the quality attribute for the reference product to ensure an adequate power for the case in which a small but sufficient number of lots are available for testing. For example, the probability of rejecting H_0 in the above two one-sided tests procedure with the equivalence margin being $\pm 1.5\sigma_R$ is 87% if the true mean difference is $0.125\sigma_R$ for a sample size of 10 test product lots and 10 reference product lots.

First, we estimate σ_R by the sample variability of the reference product, and then θ_1 and θ_2 are treated as a constant, but not a random variable in the statistical analysis.

Let X_{Tj} be the observed value of the QA of interest for Lot *j* of the test product (the proposed biosimilar product) and X_{Rj} be the observed value of the QA of interest for Lot *j* of the reference product. Since the two products are manufactured by two manufacturers, two products are independent. $\bar{X}_i = \sum_{j=1}^{n_i} X_{ij} \Big/ n_i$, and $S_i^2 = \sum_{j=1}^{n_i} \left(X_{ij} - \bar{X}_i\right) \Big/ \left(n_i - 1\right)$, where n_i is the number of lots in the i^{th} product, $i = T, R$.

Under the unequal variance of the test product and the reference product, the (1-2α)*100% CI of the mean difference in the QA of interest can be calculated as:

$$\left(\bar{X}_T - \bar{X}_R - t_\alpha(\nu)\sqrt{\frac{S_T^2}{n_T} + \frac{S_R^2}{n_R}},\ \bar{X}_T - \bar{X}_R + t_\alpha(\nu)\sqrt{\frac{S_T^2}{n_T} + \frac{S_R^2}{n_R}} \right). \quad (1)$$

where $t_\alpha(\nu)$ is the 1-α quantile and ν is the degrees of freedom calculated by Satterthwaite's approximation.

If $n_R > 1.5n_T$, the (1-2α)*100% sample size imbalanced adjusted CI of the mean difference in the QA of interest can be calculated as:

$$\left(\bar{X}_T - \bar{X}_R - t_\alpha(\nu^*)\sqrt{\frac{S_T^2}{n_T} + \frac{S_R^2}{n_R^*}},\ \bar{X}_T - \bar{X}_R + t_\alpha(\nu^*)\sqrt{\frac{S_T^2}{n_T} + \frac{S_R^2}{n_R^*}} \right). \quad (2)$$

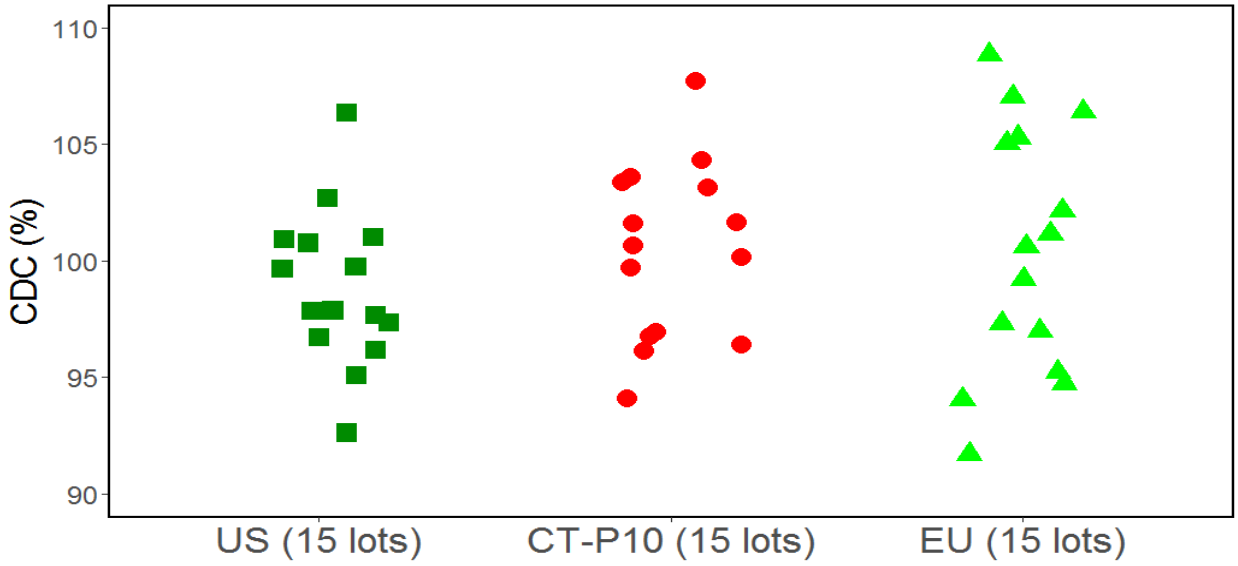
where $n_R^* = \min(n_R, 1.5n_T)$ and $\nu^* = \dfrac{\left(\frac{S_T^2}{n_T} + \frac{S_R^2}{n_R^*}\right)^2}{\frac{1}{n_T - 1}\left(\frac{S_T^2}{n_T}\right)^2 + \frac{1}{n_R - 1}\left(\frac{S_R^2}{n_R^*}\right)^2}$.

If $n_T > 1.5n_R$, we can apply a similar approach as above with $n_T^* = \min(1.5 \times n_R, n_T)$ for the CI calculation. In the following analyses, we use α=0.05.

4.3 FDA statistical equivalence testing for the relative potency of the CDC assay

The relative potency assay of the CDC assay's data points of CT-P10, US, and EU are displayed in Figure 1. CT-P10 has the largest sample mean among three products. The sample variability of CT-P10 is larger than the sample variability of US; the sample variability of CT-P10 is smaller than the sample variability of EU.

Figure 1. Scatter plot of the relative potency of the CDC assay for CT-P10, US-licensed Rituxan, and EU-approved Rituxan



Fifteen lots of all three products were included for the statistical equivalence testing for the relative potency of the CDC assay. Descriptive statistics for the relative potency of the CDC assay's data are listed in Table 2.

Table 2. Descriptive statistics for the relative potency of the CDC assay data

Product	Number of lots	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
CT-P10	15	100.47	3.83	94	108
US	15	98.93	3.28	93	106
EU	15	100.33	5.27	92	109

From Table 3, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative potency of the CDC assay pass equivalence testing.

Table 3. Equivalence testing results for the relative potency of the CDC assay

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	1.53	(-0.69, 3.75)	(-4.93, 4.93)	Yes
CT-P10 vs. EU	(15, 15)	0.13	(-2.74, 3.01)	(-7.91, 7.91)	Yes
EU vs. US	(15, 15)	1.40	(-1.35, 4.15)	(-4.93, 4.93)	Yes

4.4 FDA statistical equivalence testing for the relative potency of the ADCC assay

The applicant submitted the relative potency of the ADCC assay data at three concentration levels: 0.01 µg/mL, 0.035 µg/mL, and 0.122 µg/mL. We evaluate the data at each concentration level as follows.

4.4.1 The relative potency of the ADCC assay at 0.01 µg/mL

The relative potency of the ADCC assay's data points of CT-P10, US, and EU at 0.01 µg/mL are displayed in Figure 2. CT-P10 has the largest sample mean and sample variability among three products.

Fifteen lots of all three products were included for the statistical equivalence testing for the relative potency of the ADCC assay at 0.01 µg/mL. Descriptive statistics for the relative potency of the ADCC assay's data at 0.01 µg/mL are listed in Table 4.

From Table 5, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative potency of the ADCC assay at 0.01 µg/mL pass equivalence testing.

Figure 2. Scatter plot of the relative potency of the ADCC assay for CT-P10, US-licensed Rituxan, and EU-approved Rituxan at 0.01 µg/mL

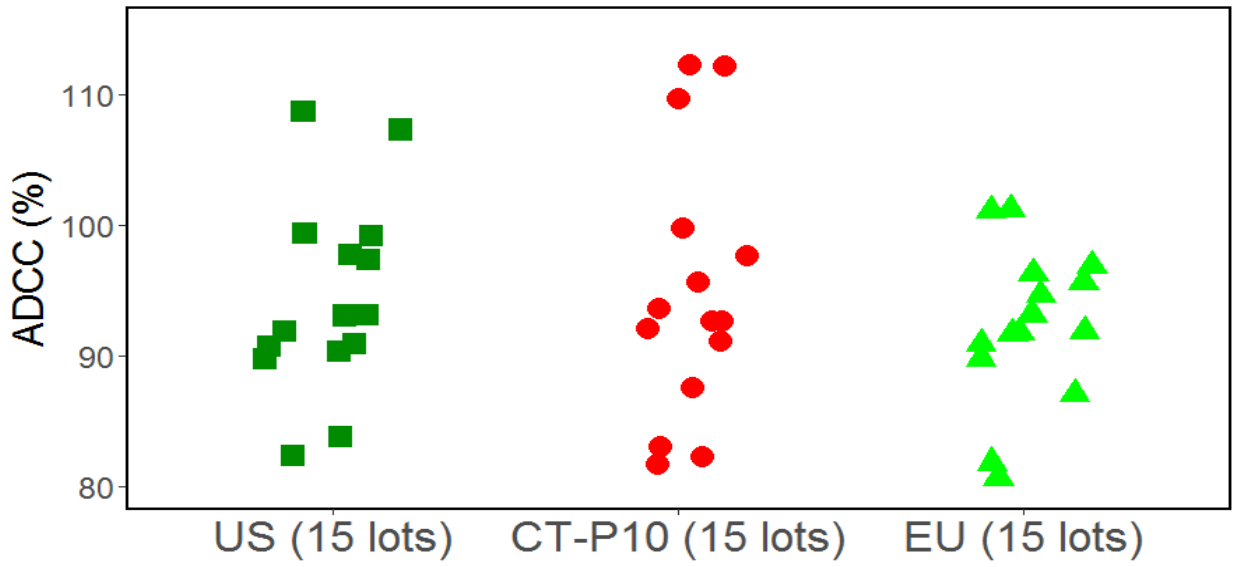


Table 4. Descriptive statistics for the relative potency of the ADCC assay data at 0.01 µg/mL

Product	Number of lots	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
CT-P10	15	95.07	10.01	82	112
US	15	94.33	7.40	82	109
EU	15	92.4	5.84	81	101

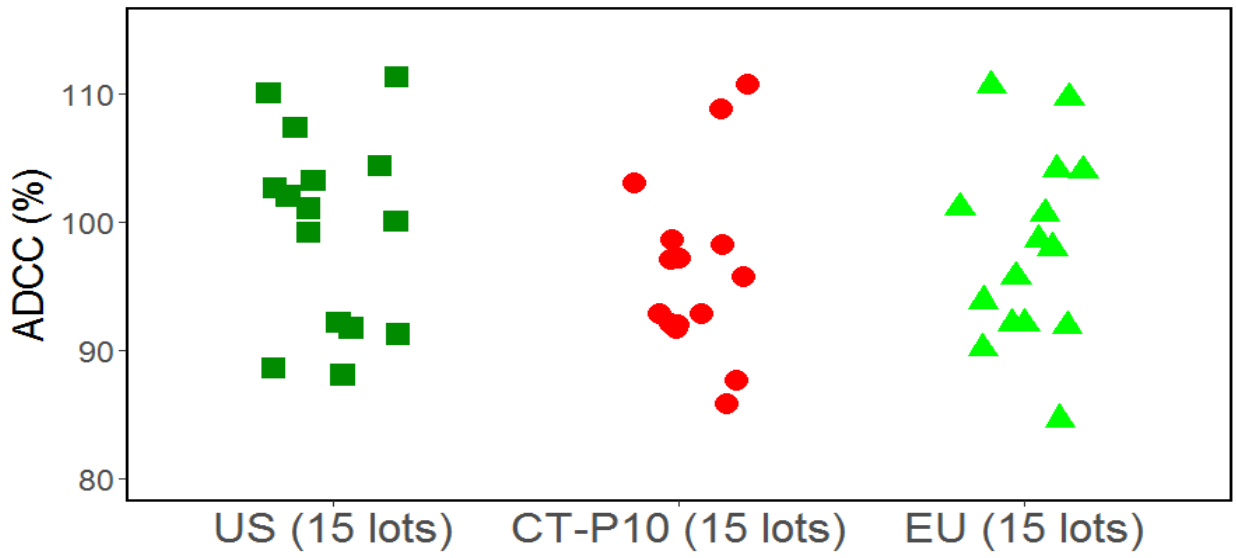
Table 5. Equivalence testing results for the relative potency of the ADCC assay at 0.01 µg/mL

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	0.73	(-4.75, 6.22)	(-11.11, 11.11)	Yes
CT-P10 vs. EU	(15, 15)	2.67	(-2.47, 7.80)	(-8.76, 8.76)	Yes
EU vs. US	(15, 15)	-1.93	(-6.08, 2.22)	(-11.11, 11.11)	Yes

4.4.2 The relative potency of the ADCC assay at 0.035 µg/mL

The relative potency of the ADCC assay's data points of CT-P10, US, and EU at 0.035 µg/mL are displayed in Figure 3. CT-P10 has the smallest sample mean and sample variability among three products.

Figure 3. Scatter plot of the relative potency of the ADCC assay for CT-P10, US-licensed Rituxan, and EU-approved Rituxan at 0.035 µg/mL



Fifteen lots of all three products were included for the statistical equivalence testing for the relative potency of the ADCC assay at 0.035 μg/mL. Descriptive statistics for the relative potency of the ADCC assay's data at 0.035 μg/mL are listed in Table 6.

Table 6. Descriptive statistics for the relative potency of the ADCC assay data at 0.035 μg/mL

Product	Number of lots	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
CT-P10	15	96.4	7.00	86	111
US	15	99.47	7.46	88	111
EU	15	97.93	7.39	85	111

From Table 7, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative potency of the ADCC assay at 0.035 μg/mL pass equivalence testing.

Table 7. Equivalence testing results for the relative potency of the ADCC assay at 0.035 μg/mL

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	-3.07	(-7.56, 1.43)	(-11.19, 11.19)	Yes
CT-P10 vs. EU	(15, 15)	-1.53	(-6.15, 3.08)	(-11.09, 11.09)	Yes
EU vs. US	(15, 15)	-1.53	(-6.00, 2.94)	(-11.19, 11.19)	Yes

4.4.3 The relative potency of the ADCC assay at 0.122 μg/mL

The relative potency of the ADCC assay's data points of CT-P10, US, and EU at 0.122 μg/mL are displayed in Figure 4. CT-P10 has the smallest sample mean among three products. In addition, CT-P10 has the largest sample variability among three products.

Fifteen lots of all three products were included for the statistical equivalence testing for the relative potency of the ADCC assay at 0.122 μg/mL. Descriptive statistics for the relative potency of the ADCC assay's data at 0.122 μg/mL are listed in Table 8.

From Table 9, the result shows that the three-way pairwise comparisons (CT-P10 vs. US, CT-P10 vs. EU, and US vs. EU) for the relative potency of the ADCC assay at 0.122 μg/mL pass equivalence testing.

Figure 4. Scatter plot of the relative potency of the ADCC assay for CT-P10, US-licensed Rituxan, and EU-approved Rituxan at 0.122 µg/mL

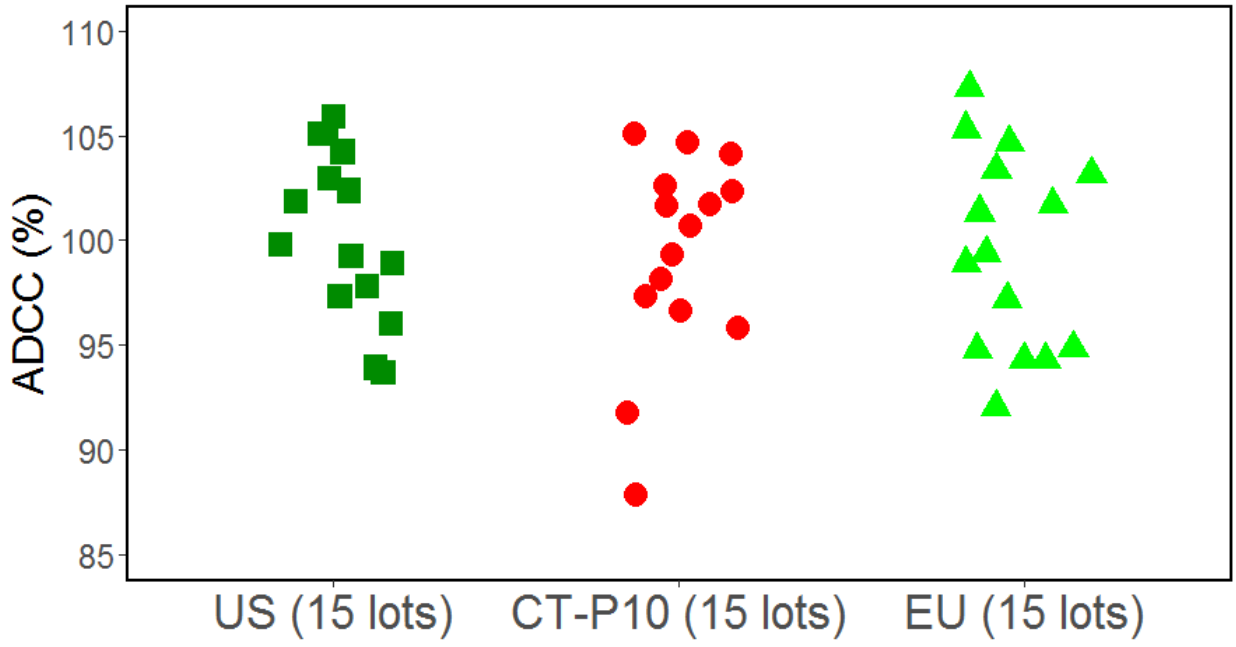


Table 8. Descriptive statistics for the relative potency of the ADCC assay data at 0.122 µg/mL

Product	Number of lots	Sample mean, mg/mL	Sample standard deviation, mg/mL	Minimum, mg/mL	Maximum, mg/mL
CT-P10	15	99.4	4.85	88	105
US	15	100.27	3.97	94	106
EU	15	99.4	4.73	92	107

Table 9. Equivalence testing results for the relative potency of the ADCC assay at 0.122 µg/mL

Comparison	Number of lots	Mean difference, %		Equivalence margin, %	Statistical Equivalence
		Estimate	90% CI		
CT-P10 vs. US	(15, 15)	-0.87	(-3.62, 1.89)	(-5.96, 5.96)	Yes
CT-P10 vs. EU	(15, 15)	0.00	(-2.98, 2.98)	(-7.10, 7.10)	Yes
EU vs. US	(15, 15)	-0.87	(-3.58, 1.85)	(-5.96, 5.96)	Yes

5 CONCLUSION AND RECOMMENDATION

The results from the statistical equivalence analyses for the relative potency of the CDC assay and the relative potency of the ADCC assay at three concentration levels supported a demonstration that the proposed biosimilar CT-P10 is highly similar to US-licensed Rituxan if the data were valid.

However, the validity of the conclusions of the statistical equivalence analyses for Tier 1 quality attributes depends on the outcomes of the pending inspection which will assess the integrity of the analytical similarity data in analytical similarity studies.

6 REFERENCE

[1] Statistical Approaches to Evaluate Analytical Similarity Guidance for Industry (2017) https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM576786.pdf

---

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

/s/

YU-TING WENG
01/17/2018

MEIYU SHEN
01/17/2018

YI TSONG
01/19/2018

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	BLA 761,088
Supplement #:	Original; 000
Drug Name:	CT-P10 (proposed biosimilar to rituximab)
Indication(s):	Biosimilar to rituximab: non-Hodgkin's lymphoma, (b) (4)
Applicant:	CELLTRION, Inc.
Date(s):	Submission date 4/28/2017; BsUFA 2/28/2018
Review Priority:	Standard.
Biometrics Division:	DBV
Statistical Reviewer:	Jingjing Ye, PhD
Concurring Reviewers:	Lei Nie, PhD, Team Leader
	Thomas Gwise, PhD, Deputy Director
Medical Division:	Division of Hematology Product (DHP), OHOP
Clinical Team:	Anand Shah, Nicole Gormley (TL)
Project Manager:	Esther Park

Keywords:

Biosimilar, Non-inferiority, Extrapolation, Advanced Follicular Lymphoma

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

CELLTRION submitted a biologics license application BLA761088 under section 351 (k) of the Public Health Service Act (PHS Act) to support CT-P10 as a biosimilar product to US-licensed Rituxan. CELLTRION is seeking licensure of CT-P10 for (b) (4) Non-Hodgkin's Lymphoma (NHL), (b) (4)

To support a demonstration of biosimilarity, a stepwise approach was used following the FDA's scientific recommendation. The stepwise approach starts with structural and functional characterization of both the proposed biosimilar product and the reference product. Results of nonclinical and/or clinical studies follow to assess remaining questions with regards to potential residual uncertainty about biosimilarity.

This review evaluated the results of CT-P10 3.3, which was a comparative randomized, multicenter, controlled, parallel-group, double-blinded study in patients with Advanced Follicular Lymphoma (AFL). AFL was not considered as sensitive population to assess biosimilarity by the agency for the oncology indication.

The primary endpoint in study CT-P10 3.3 was objective response rate (ORR) compared between CT-P10 and US-licensed Rituxan arms. The observed ORR difference between the arms was 5.7% and the 90% confidence interval was [-1.7%, 14.7%]. The lower 90% confidence interval of observed difference in ORR was greater than the pre-specified non-inferiority margin of 7%, therefore, the primary objective was met for non-inferiority of ORR. However, the 90% upper bound of the difference does not rule out non-superiority. The time-to-event secondary endpoints were not available at the data cut-off because the data were not mature.

The results of the clinical study CT-P10 3.3 did not provide support of biosimilar of CT-P10 to reference product Rituxan for the targeted oncology population in ORR. It was unclear if the difference observed in ORR between CT-P10 and Rituxan can be considered as no clinically meaningful difference in oncology population.

2 INTRODUCTION

2.1 Overview

CT-P10 is a chimeric monoclonal IgG1 subclass antibody and has been developed as a biosimilar product of the reference product, US-licensed Rituxan under Section 351(a) of the Public Health Service Act. US-licensed Rituxan (rituximab) was approved by the FDA in 1997 (BLA-103705). The proposed indications are Non-Hodgkin's Lymphoma (NHL), (b) (4)

The sponsor has conducted3 clinical studies: CT-P10 3.2 in RA indication, CT-P10 3.3 in Advanced Follicular Lymphoma (AFL) and CT-P10 3.4 in low tumor burden follicular lymphoma. In this application, the sponsor submitted data and report on CT-P10 3.2 and CT-P10 3.3 for evaluation. CT-P10 3.4 is still ongoing and the estimated completion is in 2020. For details on evaluation of CT-P10 3.2, please refer to statistical review in RA. In this review, the focus is CT-P10 3.3 clinical study in AFL.

The study CT-P10 3.3 was a phase 1/3 randomized, multicenter, controlled, parallel-group, double-blinded study in patients with AFL. The patients were randomized 1:1 to either CT-P10 or Rituxan to assess efficacy and safety. The core study period was July 28, 2014 for first patient randomized to June 27, 2016 last patient completed the core study period.

The sponsor and FDA had several communications through their development program: Type 2 meeting in Mar. 2014, July 2015, and Feb. 2016 in their development in CT-P10; Type 2 and 3 meetings in Jan. 2017 with the data generated to date; and Type 4 meeting in Apr. 2017 to obtain the feedback on the format and content of the application. In the communication and meeting minutes dated Mar. 19, 2014, FDA has recommended that the sponsor "consider a comparative clinical trial of rituximab monotherapy in the first-line treatment of patients with low-tumor burden follicular lymphoma as a more sensitive model to evaluate clinically meaningful differences between CT-P10 and rituximab." In addition, FDA commented that "if you choose to conduct a comparative clinical study in patients with AFL rather than low-tumor burden follicular lymphoma as recommended by FDA, the comparative clinical study could be insufficient to support a demonstration of no clinically meaningful differences, thus impacting the ability to support a demonstration of biosimilarity and the potential to extrapolate to other conditions of use." However, the sponsor decided submitting the CT-P10 3.3 in patients in AFL in the application for support and consideration of the CT-P10 as the biosimilar product of Rituxan in addition to CT-P10 3.2. CT-P10 3.4 is ongoing in the recommended patient population in low-tumor burden follicular lymphoma.

The study CT-P10 3.3 protocol was created on Nov. 2013 and has been amended 3 times. The last amendment was created on Mar. 16, 2016. The statistical analysis plan (SAP) was finalized on Sep. 30, 2016.

Table: List of all studies included in analysis

	Phase and Design	Treatment Period	Follow-up Period	# of Subjects per Arm	Data cut-off/Median Observation Time
CT-P10 3.3	Phase 1/3, Randomized, Parallel-group, active-controlled, double-blinded study in patients with AFL (Advanced Follicular Lymphoma)	Test product: CT-P10 (375mg/m^2 IV infusion) Reference product: Rituxan® (375mg/m^2 IV infusion) **Three Study Period** Core Study Period: up to 8 cycles (every 3 weeks) Maintenance Study Period: up to 12 cycles (every 2 months) Follow-up Period: until death or 3 years from Day 1 of Cycle 1 of the last patient (every 3 months) Co-medication (in Core Study Period): Cyclophosphamide (750 mg/m^2 IV) Vincristine (1.4 mg/m^2 [up to a maximum of 2 mg] IV) Prednisone (40 mg/m^2 oral)	2-year	CT-P10 / 70 Rituxan / 70	June 27, 2016

2.2 Data Sources

Analysis datasets, SDTM tabulations, and software codes are located on network with network path: \\CDSESUB1\evsprod\BLA761088\0000\m5\datasets

3 STATISTICAL EVALUATION

The statistical evaluation was based on data from study CT-P10 3.3.

3.1 Data and Analysis Quality

The efficacy endpoints such as overall response rate (ORR) were derived and saved in analysis datasets "ADRES", demographic and disease characteristics were derived and saved in analysis datasets "ADSL" respectively. The statistical reviewer is able to reproduce the derived ORR datasets from the BLA tabulation datasets.

3.2 Evaluation of Efficacy

3.2.1.1 Study Design

CT-P10 3.3 was a Phase 1/3, Randomized, Parallel-Group, Active-Controlled, Double-Blind Study to Demonstrate Equivalence of Pharmacokinetics and Non-inferiority of Efficacy for CT-P10 in Comparison With Rituxan, Each Administered in Combination With Cyclophosphamide, Vincristine, and Prednisone (CVP) in Patients With Advanced Follicular Lymphoma. The study diagram is summarized in Figure 1. The patients were stratified randomized by country (as recorded in the IWRS (interactive web response system)), gender, and Follicular Lymphoma International Prognostic Index (FLIPI) score (0 to 2 versus 3 to 5).

The overall objective of this study was to demonstrate similar pharmacokinetics and non-inferior efficacy of CT-P10 compared to Rituxan. With this purpose, this study was divided into 2 parts, each of which assessed 1 of 2 primary endpoints, as follows:

Part 1: The primary objective of Part 1 of the study was to demonstrate that CT-P10 is similar to Rituxan in terms of pharmacokinetics as determined by area under the serum concentration-time curve at steady state (AUC_{tau}) and maximum serum concentration at steady state ($C_{max}SS$) at Core Cycle 4.

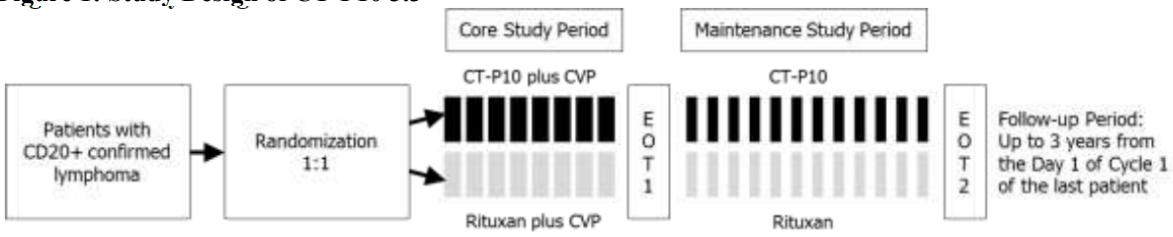
Part 2: The primary objective of Part 2 of the study was to demonstrate that CT-P10 is non-inferior to Rituxan in terms of efficacy as determined by overall response rate (ORR) (the proportion of responders who achieved complete response [CR], unconfirmed complete response [CRu], or partial response [PR]) over 8 cycles of the Core Study Period according to the 1999 International Working Group (IWG) criteria in previously untreated patients with advanced (stage III-IV) cluster of differentiation 20 positive (CD20+) follicular lymphoma (FL).

The secondary objectives of Part 2 were as follows:

- To demonstrate ORR (the proportion of responders who achieved CR or PR) over 8 cycles in the Core Study Period according to the 2007 IWG criteria.

- To evaluate additional efficacy parameters (progression-free survival [PFS], time to progression [TTP], time to treatment failure [TTF], response duration, disease-free survival [DFS], and overall survival [OS]) according to the 1999 IWG criteria and the 2007 IWG criteria for patients who underwent positron emission tomography (PET) or PET-computed tomography (PET-CT).
- To evaluate pharmacodynamics (B-lymphocyte [B-cell] kinetics), overall safety, and biomarkers of CTP10 in comparison with Rituxan.

Figure 1: Study Design of CT-P10 3.3



CVP=Cyclophosphamide, Vincristine and Prednisone; CD20=Cluster of Differentiation 20; EOT1=first End-of-Treatment Visit; EOT2=second End-of-Treatment Visit

[Source: 16.1.9 Documentation of Statistical Methods, Page 9]

CVP (cyclophosphamide (750 mg/m2 intravenous (IV)), vincristine (1.4 mg/m2 [up to a maximum of 2mg] IV), and prednisone (40 mg/m2 oral)) were co-administered along with the drug. There will be a Core Study Period of 8 cycles (each cycle of 21 ± 3 days) in which the CT-P10 or Rituxan (375 mg/m² IV) and CVP will be administered. Up to 8 treatment cycles of R-CVP and CT-P10 will be administered after which patients will be evaluated for response to the treatment. CT-P10 or Rituxan (375 mg/m2 IV) will be administered alone as a maintenance therapy in patients who have a response after Cycle 8 of the Core Study Period (as assessed at the first EOT Visit). The Maintenance Study Period will be continued up to maximum of 12 cycles, with CT-P10 or Rituxan administered every 2 months for a maximum of 2 years and those who are not eligible for the Maintenance Study Period or complete the Maintenance Study Period will enter the Follow-up Period.

The study enrolled 140 patients, who were randomly assigned to treatment for Part 2 of the study, including 159 screened patients and 121 patients randomly assigned to treatment for Part 1.

The sample size was calculated based on the proportion of patients with either a CR or CRu, PR response (the "responders"). An expected ORR of 81% has been assumed for MabThera compared to CVP alone based on Marcus et al 2005 study. A 7% non-inferiority margin has been pre-specified for the trial. In addition, 13% dropout rate was assumed with 80% power.

Statistical reviewer's comment: To show biosimilar between the study drug and reference drug, comparative clinical studies should be designed to establish evidence that "the proposed product is neither inferior…nor superior to the reference product by more than a (possibly different) specified margin" as recommended by FDA's guidance. Instead of designing as an equivalence study, the sponsor designed a non-inferiority study. One important difference between the equivalence and non-inferiority study is that equivalence study would control both the upper and

lower bounds of the difference in the treatment effect, while non-inferiority is designed to control only the lower bound.

The study was under-powered for the primary endpoint ORR to show non-inferiority. Based on the statistical reviewer's calculation, the power was < 30% for 5% type-I error rate given the 7% non-inferiority margin and the assumed ORR of 81%. The required sample size for ORR 81% and non-inferiority margin 10% would need 532 patients in the study. The sample size calculation was summarized in the table below given assumptions of 5% type-I error rate and 80% power:

ORR	***Margin***	***Sample Size***
81%	*7%*	*1078*
81%	*10%*	*532*
90%	*7%*	*646*
90%	*10%*	*322*

Reference: FDA guidance on Scientific Considerations in Demonstrating Biosimilarity to a Reference Product, 2015, https://www.fda.gov/downloads/drugs/guidances/ucm291128.pdf

3.2.1.2 Efficacy Endpoints

The study design has 2 primary endpoints of PK similarity and non-inferiority of efficacy.

The primary efficacy endpoint was the ORR (CR + CRu +PR) during the Core Study Period as per the 1999 IWG criteria. The ORR was defined as the proportion of patients with a best overall response of CR, CRu or PR. The best overall response was the best response recorded from randomization until progressive disease, start of new anticancer therapy, end of Core Study Period or death, whichever came first. If best overall responses of two or more assessment visits were same, then the best overall response of earlier assessment date were used.

For secondary endpoints:

The secondary efficacy endpoint on tumor response evaluation was the ORR (CR + PR) during the Core Study Period as per 2007 IWG criteria for patients who underwent PET or PET-CT. The best overall response were applied as specified in earlier section, except for CRu. The proportion of overall responders over 8 cycles of treatment were summarized by treatment group.

For time-to-event endpoints, PFS, TTP, TTF, response duration, DFS, OS and follow-up duration endpoints will be performed. The follow-up period would be up to 3 years from the day 1 of cycle 1 of the last patient.

PFS is defined as the interval between randomization and disease progression, relapse after complete remission or death from any cause, whichever occurs first (Cheson et al 2007).

Patients would be censored without disease progression/relapse or death at the time of the last tumor assessment before starting another anticancer therapy.

Response duration is the time to first documentation of relapse or progression from the first time when criteria of response (CR, CRu or PR) is met (Cheson et al 1999; Cheson et al 2007).

DFS was measured from the time of occurrence of attained first CR or CRu to disease recurrence or death as a result of lymphoma or acute toxicity of treatment, whichever occurs first (Cheson et al 1999; Cheson et al 2007).

OS is defined as the interval between randomisation and death from any cause (Cheson et al 2007).

3.2.2 Statistical Methodologies

The proportion of responders (CR + CRu + PR) over 8 cycles of RCVP treatment was compared between CT-P10 and Rituxan. If the true difference between CT-P10 and Rituxan should lie entirely on the positive side of the non-inferiority margin of -7%, then non-inferiority of CT-P10 to Rituxan with respect to response rates will be claimed for this primary endpoint.

A time-to-event analysis by the Kaplan Meier method will be performed for each of the PFS, TTP, TTF, response duration, DFS, OS and follow-up duration endpoints in the ITT population according to the 1999 IWG criteria and 2007 IWG criteria for patients who underwent positron emission tomography (PET) or PET-computed tomography (PETCT), which applicable. The number of events and number of censored subjects by treatment group will be summarized with censoring reasons. The survival rates per 6 months will be estimated by the Kaplan Meier method and presented along with their corresponding 95% CI.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

This study was conducted in 64 study centers in Europe, Africa, Asia Pacific, and Latin America. Of these study centers, there were 184 screened patients and 140 patients randomly assigned to treatment for Part 2 of the study including 159 screened patients and 121 patients randomly assigned to treatment for Part 1.

Analysis population

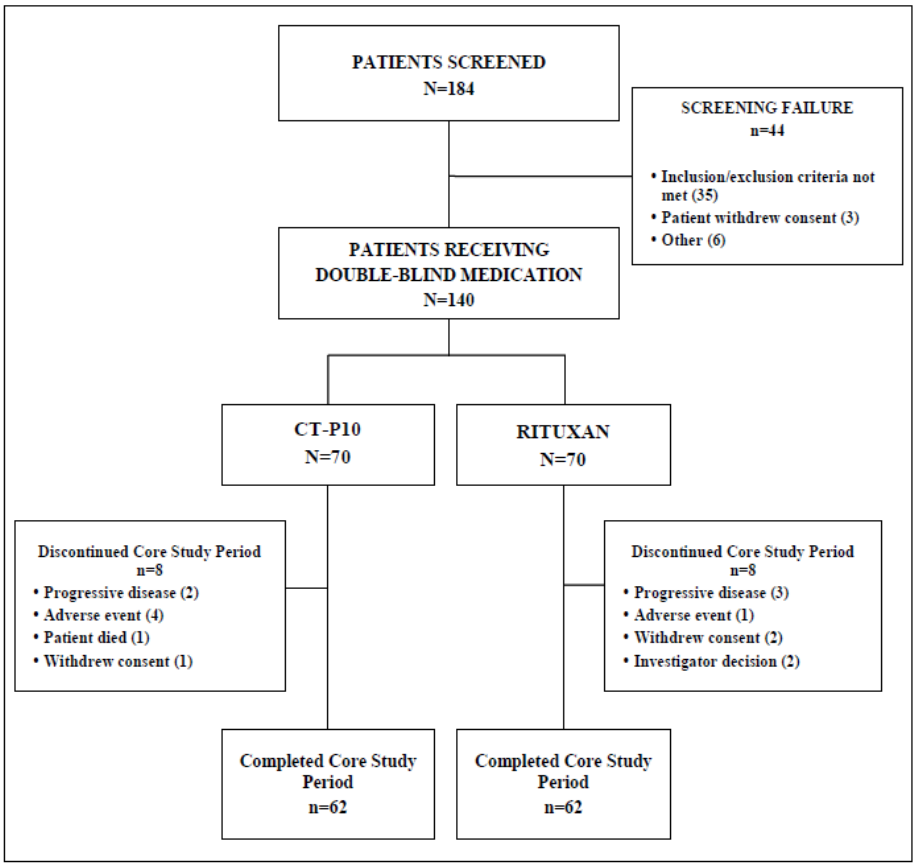
The ITT (Intent-to-Treat) population was defined as all patients enrolled and randomly assigned to receive a dose of study drug, regardless of whether or not any study treatment dosing was completed. Patients in the ITT population were analyzed according to the treatment they received based on the “as-treated” principle.

The safety population was defined as all randomly assigned patients who received at least one dose (full or partial) of study drug (CT-P10 or Rituxan).

Subject disposition

A total of 140 patients were randomly 1:1 assigned to CT-P10 or Rituxan. The subject disposition is provided in Figure 2 and Table 1. It appears to be balanced for early discontinuation between the arms, although the adverse events were 4 in CT-P10 arm compared to 1 in Rituxan arm. However, the number of patients was too small to draw any conclusions.

Figure 2: CT-P10 3.3 Subject Disposition



Source: Post-text Table 14.1.1.

[Source: CT-P10 3.3 Clinical Study Report Page 121, Figure 10-2]

Table 1: Subject disposition ITT population in CT-P10 3.3

	CT-P10 (N=70)	Rituxan (N=70)	Total (N=140)
	Number (%) of Patients		
Total number of patients in Part 2			
Screened[1]			184
Primary reason for screening failure[2]			
Inclusion/exclusion criteria not met			35
Patient withdrew consent			3
Other			6
Randomized	70 (100.0)	70 (100.0)	140 (100.0)
Initiated core study treatment[3]	70 (100.0)	70 (100.0)	140 (100.0)
Completed the Core Study Period	62 (88.6)	62 (88.6)	124 (88.6)
Early discontinued before completion of the Core Study Period	8 (11.4)	8 (11.4)	16 (11.4)
Primary reason for discontinuation			
Progressive disease[4]	2 (2.9)	3 (4.3)	5 (3.6)
Adverse event[5]	4 (5.7)	1 (1.4)	5 (3.6)
Patient died (tumour lysis syndrome)[6]	1 (1.4)	0	1 (0.7)
Withdrew consent	1 (1.4)	2 (2.9)	3 (2.1)
Investigator decision[7]	0	2 (2.9)	2 (1.4)
Number of deaths	3 (4.3)	0	3 (2.1)
Reason for death			
Lymphoma[8]	2 (2.9)	0	2 (1.4)
Adverse event (tumor lysis syndrome)[6]	1 (1.4)	0	1 (0.7)

Abbreviation: ITT, intent-to-treat.

Note: Unless otherwise stated, percentages were calculated using the intent-to-treat population as the denominator.

1. Includes screening failures and randomly assigned patients. If a patient was screened and randomly assigned, the treatment assigned was displayed in the "Randomized" row.
2. This summary includes screening failures and nonrandomized patients only.
3. Includes patients who were randomly assigned and initiated treatment.
4. Last administered dose was Core Cycle 2 for Patient 1201-3002 and Core Cycle 4 for Patient 2805-3001 in the CT-P10 treatment group and Core Cycle 2 for Patient 2501-3003, Core Cycle 4 for Patient 3205-3006 and Core Cycle 7 for Patient 5106-3002 in the Rituxan treatment group.

[Source: CT-P10 3.3 Clinical Study Report Page 122, Table 10-2 and statistical reviewer's analysis]

Subject demographics and baseline disease characteristics

Baseline disease characteristics and demographics were summarized in Table 2 and Table 3. The baseline and disease characteristics appear to be balanced between the arms.

Table 2: CT-P10 3.3 Demographic and Baseline Characteristics

Parameter	CT-P10 (N=70)	Rituxan (N=70)	Total (N=140)
Statistics	Number (%) of Patients		
Age (years)			
n	70	70	140
Mean (SD)	55.0 (13.76)	56.5 (14.12)	55.8 (13.91)
Median	57.0	58.5	57.5
Minimum, maximum	30, 85	26, 84	26, 85
Gender			
Male	30 (42.9)	33 (47.1)	63 (45.0)
Female	40 (57.1)	37 (52.9)	77 (55.0)
Height (cm)			
n	70	70	140
Mean (SD)	166.9 (9.19)	164.9 (9.83)	165.9 (9.53)
Median	166.0	165.0	165.5
Minimum, maximum	142, 188	145, 188	142, 188
Weight (kg)			
n	70	70	140
Mean (SD)	73.17 (15.669)	72.00 (15.317)	72.59 (15.449)
Median	71.00	71.25	71.00
Minimum, maximum	44.6, 130.0	43.7, 105.0	43.7, 130.0
Race			
White or Caucasian	51 (72.9)	52 (74.3)	103 (73.6)
Asian	11 (15.7)	13 (18.6)	24 (17.1)
Hispanic or Latino	6 (8.6)	3 (4.3)	9 (6.4)
Black or African American	2 (2.9)	0	2 (1.4)
Other	0	1 (1.4)	1 (0.7)
Not allowed by investigator country regulation	0	1 (1.4)	1 (0.7)
Region			
European Union	28 (40.0)	32 (45.7)	60 (42.9)
Non-European Union	42 (60.0)	38 (54.3)	80 (57.1)

Parameter	CT-P10 (N=70)	Rituxan (N=70)	Total (N=140)
Statistics	Number (%) of Patients		
ECOG performance status at screening			
0	44 (62.9)	47 (67.1)	91 (65.0)
1	25 (35.7)	22 (31.4)	47 (33.6)
2	1 (1.4)	1 (1.4)	2 (1.4)
3	0	0	0
4	0	0	0
5	0	0	0

Abbreviations: ECOG, Eastern Cooperative Oncology Group; ITT, intent-to-treat; SD, standard deviation.
Source: Post-text Table 14.1.4.

[Source: CT-P10 3.3 Clinical Study Report Page 129-130, Table 11-4 and statistical reviewer's analysis]

Table 3: CT-P10 3.3 Baseline Ann Arbor Principal Staging and FLIPI Score

Parameter	CT-P10 (N=70)	Rituxan (N=70)	Total (N=140)
Result	Number (%) of Patients		
Ann Arbor principal staging at screening			
Stage I	0	0	0
Stage II	0	0	0
Stage III	21 (30.0)	36 (51.4)	57 (40.7)
Stage IV	49 (70.0)	34 (48.6)	83 (59.3)
FLIPI score at screening			
0	0	0	0
1	8 (11.4)	6 (8.6)	14 (10.0)
2	25 (35.7)	21 (30.0)	46 (32.9)
3	23 (32.9)	30 (42.9)	53 (37.9)
4	10 (14.3)	12 (17.1)	22 (15.7)
5	4 (5.7)	1 (1.4)	5 (3.6)

Abbreviations: FLIPI, Follicular Lymphoma International Prognostic Index; ITT, intent-to-treat.

Source: Post-text Table 14.1.6.

[Source: CT-P10 3.3 Clinical Study Report Page 133, Table 11-6 and statistical reviewer's analysis]

Protocol deviation

A major protocol deviation was one that might affect the interpretation of study results or the patient's rights, safety, or welfare. A summary was provided in Table 4.

Table 4: Major protocol deviations in ITT population

	CT-P10 (N=70)	Rituxan (N=70)	Total (N=140)	Excluded Populations[1]
	Number (%) of Patients			
Major protocol deviations				
Non-compliance with I/E criteria	3 (4.3)	0	3 (2.1)	PP
Other reasons for exclusion				
No response evaluation after treatment	1 (1.4)	2 (2.9)	3 (2.1)	PP

Abbreviations: I/E, inclusion/exclusion; ITT, intent-to-treat; PP, per-protocol.

1. Indication of the populations of analysis from which the protocol deviation excluded patients.

Source: Post-text Table 14.1.3.

[Source: CT-P10 3.3 Clinical Study Report Page 124, Table 10-3 and statistical reviewer's analysis]

3.2.4 Results and Conclusions

3.2.4.1 Primary Analysis Results

The results of the primary endpoint of ORR based on central review were summarized in Table 5 and Table 6.

Table 5: Primary analysis on ORR (Central Review)

Population	CT-P10	Rituxan	Difference[1]
Overall Response	Number (%) of Patients		%

ITT population	n=70	n=70	
ORR (CR+CRu+PR)	67 (95.7)	63 (90.0)	5.7
CR	21 (30.0)	15 (21.4)	
CRu	6 (8.6)	8 (11.4)	
PR	40 (57.1)	40 (57.1)	
Stable disease	1 (1.4)	2 (2.9)	
RD/PD	1 (1.4)	2 (2.9)	
Unable to assess[2]	0	1 (1.4)	
Missing[3]	1 (1.4)	2 (2.9)	

Abbreviations: CR, complete response; CRu, unconfirmed complete response; ITT, intent-to-treat; PP, per-protocol; IWG, International Working Group; ORR, overall response rate; PR, partial response; RD/PD, relapsed disease/progressive disease.

[Source: CT-P10 3.3 Clinical Study Report Page 148, Table 11-12 and statistical reviewer's analysis]

Table 6: FDA's Calculation of ORR and the difference between ORR

	CT-P10, N=70	Rituximab, N=70	Diff [90% CI]
ORR	95 7 [88.1%, 98.5%]	90% [80.8%, 95.1%]	5.7% [-1.7%, 14.7%]

[Source: Statistical reviewer's analysis]

Statistical reviewer's comment: The study was not designed to show equivalence. The study was designed to show non-inferiority of CT-P10 compared to Rituxan even though the study was under-powered for the pre-specified non-inferiority margin 7%.

The study met the non-inferiority margin of 7% based on the results, however, in FDA Biosimilar Guidance (2015), it said "…should be neither inferior… nor superior to the reference product". The study showed the 90% confidence interval between -1.7% and 14.7%. The upper bound of the confidence interval implies that we cannot rule out superiority.

In addition, in Marcus 2005 study, the ORR was 81% and that is the basis of the sample size calculation. In this study, the Rituxan ORR was at 90%, therefore, the ORR observed was higher than Marcus 2005 study. As a result, the constancy assumption of the historical studies was in question.

Reference: FDA guidance on Scientific Considerations in Demonstrating Biosimilarity to a Reference Product, 2015, https://www.fda.gov/downloads/drugs/guidances/ucm291128.pdf

3.2.4.2 Additional Analysis on Primary Endpoint

Because the current study observed a higher ORR rate than the historical Marcus 2005 study, to further examine the R-CVP ORR over the years, Table 7 gave the historical studies in the literature on the R-CVP ORR rates. Assume the historical control groups are exchangeable with the current control group, the statistical reviewer ran fixed-effect and random-effect meta-analysis to report pooled ORR and the 95% confidence interval from the historical studies. The results were given in Table 8. The results were very similar between fixed and random-effect meta-analysis. Even though the observed ORR in R-CVP arm increased over the years, the upper

bound of the 95% confidence interval based on meta-analysis excluded 90% ORR rates observed in the current trial. This again raised question of constancy assumption.

Table 7: Historical studies on ORR

Trial	Author (year)	R-CVP		CVP	
		N	ORR* (95% CI)	N	ORR* (95% CI)
1	Marcus (2005)	162	81% (74-87%)	159	57% (49-65%)
2	Kalinka (2005)			10	62% (31-92%)
3	Hagenbeek (2006)			187	52% (44-59%)
4	Levitt (2009)	7	100%		
5	Moccia (2010)	251	86% (81-90%)		
6	Salles (2011)	272	82% (77-86%)		
7	Nastoupil (2011)	247	88% (83-92%)		
8	Federico (2013)	168	88% (83-92%)		

[Source: Literature]

Table 8: FDA's meta-analysis on ORR on R-CVP using historical studies

FDA Analysis	% [95% CI]
Fixed Effect	85.4 [83.4, 87.5]
Random Effect	85.3 [82.5, 88.0]

[Source: statistical reviewer's analysis]

To further examine the baseline characteristics between Marcs (2005) trial and the current CT-P10 3.3 trial, the statistical reviewer also compared the two studies. The results were summarized in Table 9. As can be seen from the table, the baseline characteristics seem comparable between Marcus 2005 study and the current study. The increased ORR rate in the current study was not explained by the baseline characteristics.

Table 9: Baseline characteristics comparison between CT-P10 3.3 trial and Marcus 2005 Study

Characteristics	CVP,#(%)	R-CVP,#(%)	CT-P10 3.3
No.	159	162	140
Age			
<40	16 (10)	24 (15)	20 (14)
40-50	45 (28)	48 (30)	26 (19)
51-60	54 (34)	49 (30)	34 (24)
>60	44 (28)	41 (25)	60 (43)

Median	53	52	57.5
Male	85 (54)	88 (54)	63 (45)
Performance Status (ECOG)			
0	90 (57)	93 (57)	91 (65)
1	60 (38)	65 (40)	47 (34)
>1	8 (5)	4 (3)	2 (1)
Not evaluable/missing	1 (1)	0	0
Stage (Ann Arbor)			
II	2 (1)	2 (1)	0
III-IV	157 (99)	159 (98)	140 (100)
Not evaluable/missing	0	1 (1)	0
FL Prognostic Index Score			
0-2	75 (47)	80 (49)	60 (43)
3-5	75 (47)	71 (44)	80 (57)
Not evaluable/missing	9 (6)	11 (7)	0
Elevated LDH	39 (26)	39 (26)	39 (28)

[Source: Statistical reviewer's analysis and literature by Marcus 2005 study]

3.2.4.3 Secondary Analysis Results

The secondary endpoints including PFS, TTP, TTF, response duration, DFS, OS and follow-up duration were not included in the original submission and the sponsor indicated that the information will be included in the final report.

An information request was sent to the sponsor in Sep. 2017 to request the up-to-date secondary time-to-event endpoints. The sponsor responded using the data cut-off date on Jan. 31, 2017 and summarized the median time-to-event endpoints in Table 10. The medians were not reached in either treatment arms.

Table 10: Time-to-event endpoints using not mature data

	CT-P10 (N=70)	Rituxan (N=70)
PFS (95% CI)	NE [NE, NE]	NE [19.4, NE]

DoR (95% CI)	NE [NE, NE]	NE [16.9, NE]
OS (95% CI)	NE [NE, NE]	NE [NE, NE]

3.3 Evaluation of Safety

Please refer to clinical review of this application for safety results and conclusions for safety.

3.4 Benefit-Risk Assessment

Whether the submission demonstrated an overall favorable risk-benefit profile on CT-P10 over Rituximab is deferred to the clinical team reviewing this submission.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

The sponsor did not report any subgroup analysis. The statistical review analyzed the subgroup for primary endpoint ORR. The results were summarized in Table 11. No outlier groups were identified.

Table 11: Subgroup evaluation of ORR between the arms

Subgroup		CT-P10	Rituxan	Diff (90% CI)
Gender	Female	97.5% [87.1, 99.9]	97.3% [86.2, 99.9]	0.2% [-9.1, 11.3]
	Male	93.3% [78.7, 98.2]	81.8% [65.6, 91.4]	11.5% [-3%, 26.6%]
Race	White	96.1% [86.8, 98.9]	92.3% [81.8, 97]	3.8% [-4.7, 14.2]
	Non-white	94.7% [75.4, 99.7]	83.3% [60.8, 94.2]	11.4% [-7.4, 32]
Age	>60	96.6% [82.8, 99.8]	87.1% [71.1, 94.9]	9.5% [-5.8, 25.9]
	<=60	95.1% [83.9, 98.7]	92.3% [80, 97.3]	2.8% [-9.1, 14.8]
Geographic Region	European union	100% [87.9, 100]	90.6% [75.8, 96.8]	9.4% [-0.5, 24.7]
	Non-European union	92.9% [81, 97.4]	89.5% [75.9, 95.8]	3.4% [-8.6, 18.1]

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

As communicated by FDA in Mar. 2014, CT-P10 3.3 study was done in AFL population. FDA recommended the study to conduct in low-tumor burden follicular lymphoma rather than AFL because low-tumor burden follicular lymphoma is considered a more sensitive population to demonstrate biosimilarity in oncology population. That population is ongoing in CT-P10 3.4 study. The results from CT-P10 3.4 would be considered more scientifically sufficient in providing evidence of demonstrating safety, purity, and potency in an appropriate condition of use.

Even though the study was underpowered to show non-inferiority with the margin of 7% under the assumption that the ORR rate is the same from both arms (please check. If assuming ORR is sufficiently larger in the biosimilar arm, the study will not be under powered even for superiority), the observed ORR met the pre-specified non-inferiority criteria. However, the 90% upper bound of the difference is 14.7%. Thus we cannot rule out non-superiority with high confidence.

It seems the patient baseline demographic was similar between Marcus 2005 study and CT-P10 3.3. But the ORR was observed to be higher in the current study than Marcus 2005 study.

5.2 Conclusions and Recommendations

The results of the clinical study CT-P10 3.3 alone do not provide sufficient support of biosimilar of CT-P10 to reference product Rituxanin oncology population. It was unclear if the difference observed in ORR between CT-P10 and Rituxan can be considered as no clinically meaningful difference in oncology population.

Reference

1. FDA guidance on Scientific Considerations in Demonstrating Biosimilarity to a Reference Product, 2015, https://www.fda.gov/downloads/drugs/guidances/ucm291128.pdf

2. [Marcus R, Imrie K, et al (2005)] CVP chemotherapy plus rituximab compared with CVP as first-line treatment for advanced follicular lymphoma. *Blood*; 105:1417–1423.

3. [Kalinka E, Juszczyński P, et al (2005)] The role of HLA DRB1 genetic polymorphisms in non-Hodgkin's lymphomas. *Przegl Lek.* 2005;62(12):1422-6. Polish.

4. [Hagenbeek A, Eghbali H, Monfardini S, et al (2006)] Phase III intergroup study of fludarabine phosphate compared with cyclophosphamide, vincristine, and prednisone

chemotherapy in newly diagnosed patients with stage III and IV low-grade malignant non-Hodgkin's lymphoma. *J Clin Oncol*. 2006 Apr 1;24(10):1590-6.

5. [Levitt MJ, Gharibo M, et al (2009)] Accelerated R-COP: a pilot study for the treatment of advanced low grade lymphomas that has a high complete response rate. *J Chemother*. 2009 Aug; 21(4):434-8.

6. [Moccia, A, Hoskins, P et al (2010)] Front-line therapy with rituximab, cyclophosphamide, vincristine, and prednisone (R-CVP) followed by 2 years of rituximab maintenance for follicular lymphoma (FL) is associated with excellent outcomes and improved progression-free survival (PFS) in comparison to no maintenance. *Blood (ASH Annual Meeting Abstracts)* 2010 116: Abstract 1803.

7. [Salles G, Seymour J et al (2011)] Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomized controlled trial. *Lancet* 2011 Jan 1; 377(9759):42-51.

8. [Nastoupil L, Sinha R et al (2011)] A comparison of the effectiveness of first-line chemoimmunotherapy regimens for follicular lymphoma (FL) used in the United States. *ASH* 2010, Abstract 97.

9. [Federico M, Luminari S et al (2013)] R-CVP Versus R-CHOP Versus R-FM for the Initial Treatment of Patients With Advanced-Stage Follicular Lymphoma: Results of the FOLL05 Trial Conducted by the Fondazione Italiana Linfomi. *J Clin Oncol*. 2013 Apr 20;31(12):1506-13.

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

/s/

JINGJING YE
01/11/2018

LEI NIE
01/11/2018

THOMAS E GWISE
01/11/2018