CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

761073Orig1s000

STATISTICAL REVIEW(S)

US 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

Biometrics Division: VI

BLA NO.:	761073
DATE RECEIVED BY OB:	Jan 17, 2019
DRUG NAME:	ABP 980 (proposed biosimilar to Herceptin) Kanjinti® (trastuzumab-anns)
DOSAGE FORM:	lyophilized powder in a strength of 420 mg/vial
INDICATION:	HER2 overexpressing breast cancer
APPLICANT:	Amgen
REVIEW FINISHED:	May 24, 2019
STATISTICAL REVIEWER:	Yu-Yi Hsu
SECONDARY REVIEWER:	Meiyu Shen
PROJECT MANAGER:	Charlene Wheeler

Concur: Thomas Permutt, Ph.D., Associate Director, CDER/OTS/OB

Distribution:

Yi Tsong, Ph.D., Division Director, OMPT/CDER/OTS/OB/DB VI

Charlene Wheeler, Senior Regulatory Health PR, OMPT/CDER/OND/OHOP/DOPI

Lei Zhang, Ph.D, OMPT/CDER/OPQ/OBP/DBRRII

Table of Contents

List of Tables

List of Figures

1 Executive summary and recommendation

To demonstrate analytical similarity, three critical quality attributes (QAs): proliferation inhibition bioassay (potency), antibody-dependent cellular cytotoxicity activity (ADCC) using an engineered NK92 cell line (ADCC-NK92) and ADCC activity using a peripheral blood mononuclear cell (ADCC-PBMC) were assigned as Tier-1 in this resubmission. Tier 1 statistical equivalence testing was conducted using equivalence margins of ±1.5 times of the standard deviation of the reference product, estimated by $\hat{\sigma}_R$, where the subscript R represents the comparator product.

Samples from 27 lots of ABP 980 (Amgen), 29 lots of US-licensed Herceptin, and 33 lots of Herceptin sourced at EU were used for evaluating the similarity of tier-1 QAs using the equivalence testing. The results are summarized in Table 1 to Table 3.

Table 1 Results of equivalence testing for proliferation inhibition bioassay (potency)

Test Product (Number of lots)	Comparator (Number of lots)	Mean Difference, %		Equivalence Margin, %	Conclusion
		Estimate	90% CI		
ABP 980 (27)	US (29)	2.19	(-2.91, 6.39)	(-14.61, 14.61)	Pass
ABP 980 (27)	EU (33)	4.23	(-1.59, 8.23)	(-15.11, 15.11)	Pass
EU (33)	US (29)	-2.04	(-6.58, 3.08)	(-14.61, 14.61)	Pass

Table 2 Equivalence testing results for ADCC-NK92 using the MWCMLE for parameter margin

Test Product (Number of lots)	Comparator (Number of lots)	Mean Difference, %		Equivalence Margin, %	Conclusion
		Estimate	90% CI		
ABP 980 (27)	US (29)	25.12	(7.70, 33.92)	(-34.36, 34.36)	Pass
ABP 980 (27)	EU (33)	19.83	(3.10, 28.27)	(-38.08, 38.08)	Pass
EU (33)	US (29)	5.29	(-7.17, 16.27)	(-34.36, 34.36)	Pass

Table 3 Equivalence testing results for ADCC-PBMC using the MWCMLE for parameter margin

Test Product (Number of lots)	Comparator (Number of lots)	Mean Difference, %		Equivalence Margin, %	Conclusion
		Estimate	90% CI		
ABP 980 (25)	US (27)	15.72	(1.41, 23.85)	(-31.48, 31.48)	Pass
ABP 980 (25)	EU (33)	14.54	(-0.18, 22.50)	(-33.92, 33.92)	Pass
EU (33)	US (27)	1.18	(-9.54, 11.58)	(-31.48, 31.48)	Pass

The results from the statistical equivalence analyses of the proliferation inhibition, ADCC-NK92 and ADCC-PBMC data support a demonstration that ABP 980 and US-Herceptin are highly similar and support the analytical portion of the scientific bridge to justify the relevance of EU-Herceptin data from the comparative clinical study.

2 Overview

ABP 980 is a proposed biosimilar to the US-licensed Herceptin (trastuzumab) submitted under Section 351(k) of the Public Health Service Act. The applicant submitted analytical similarity assessments using a three-way bridge approach to demonstrate analytical similarity between ABP 980 and the US-Herceptin and to establish a scientific bridge between the US-licensed Herceptin and EU-approved Herceptin.

Three critical quality attributes, proliferation inhibition bioassay (potency), antibody-dependent cellular cytotoxicity activity (ADCC) using an engineered NK92 cell line (ADCC-NK92) and ADCC activity using a peripheral blood mononuclear cell (ADCC-PBMC) were assigned as Tier-1 in this resubmission and were tested using samples from multiple biosimilar, US-Herceptin (440 mg presentation) and EU-Herceptin (150 mg presentation) lots.

Comments reagarding the applicant's statistical equivalence testing for proliferation inhibition, ADCC-NK92 and ADCC-PBMC are provided in Section 3. Independent statistical analyses and descriptions of received data are summarized in Section 4. Conclusion is provided in Section 5.

3 Applicant's statistical equivalence testing

In this submission, the applicant conducted 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 comparator product lots, for proliferation inhibition, ADCC-NK92, and ADCC-PBMC.

The applicant observed a shift in ADCC-NK92 activity for US-Herceptin and EU-Herceptin lots with expiration date beyond July 2018. The applicant defined US-Herceptin and EU-Herceptin with expiration date beyond July 2018 as post-shift lots. The applicant claim that the shift caused lower sample mean of ADCC-NK92 for US-Herceptin. For that reason, the applicant applied both the mean based and the median based methods for ADCC-NK92 and claim that the median based method is less affected by the shift. However, it is not clear if using the multiplier 1.5 combing with IQR as the equivalence margin is justifiable. Furthermore, comparison in medians provides different interpretation from the comparison of means in terms of central tendency.

The CMC statistical reviewer performs an independent statistical analysis in the next section. We use all available data that includes data from all available lots for tests of Tier-1 QAs: proliferation inhibition, ADCC-NK92 and ADCC-PBMC.

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. This review focuses on the Tier 1 statistical equivalence testing

4.1 Data analyzed

The applicant tested multiple lots of ABP 980, US- Herceptin and EU- Herceptin for Tier-1 QAs. The numbers of lots tested for Tier-1 QAs are listed in Table 4.

Table 4 Number of lots provided for Tier-1 quality attributes

Quality attribute	ABP 980	US-Herceptin	EU-Herceptin
Proliferation inhibition	27	29	33
ADCC activity using NK92	27	29	33
ADCC activity using PBMC	25	27	33

The applicant provided data from 27 biosimilar lots, including 7 drug product lots with 150mg presentation (DP-150mg), 6 drug product lots with 440mg presentation (DP-440mg) and 14 drug substance (DS) lots for assessing analytical similarity.

Table 5 shows descriptive statistics for the proliferation inhibition bioassay data grouped by lot type (DS or DP). Table 6 shows descriptive statistics for ADCC using NK cell for ABP 980 lots grouped by lot types.

Table 7 shows descriptive statistics for ADCC using PBMC.

Table 5 Descriptive statistics for the proliferation inhibition bioassay of ABP 980 grouped by type (DS or DP)

Subgroup	Number of lots	Mean, %	Sample standard deviation, %	Min, %	Max, %
DP – 150mg	7	109.02	7.75	96	118
DP – 440mg	6	107.50	3.16	103	110
DS	14	113.52	5.83	105	123
All (DP+DS)	27	111.01	6.33	96	123

Table 6 Descriptive statistics for the ADCC-NK92 of ABP 980 grouped by type (DS or DP)

Subgroup	Number of lots	Mean, %	Sample standard deviation, %	Min, %	Max, %
DP – 150mg	7	97.83	5.70	88	103
DP – 440mg	6	103.84	13.9	79	117
DS	14	104.32	6.59	93	117
All (DP+DS)	27	102.53	8.63	79	117

Table 7 Descriptive statistics for the ADCC-PBMC of ABP 980 grouped by type (DS or DP)

Subgroup	Number of lots	Mean, %	Sample standard deviation, %	Min, %	Max, %
DP – 150mg	8	95.46	11.02	82	115
DP – 440mg	7	98.54	11.69	85	115
DS	11	99.45	10.30	89	124
All (DP+DS)	25	97.30	10.22	82	124

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 (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):

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

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:

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

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.

$$(L, U) =$$

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

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

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

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].

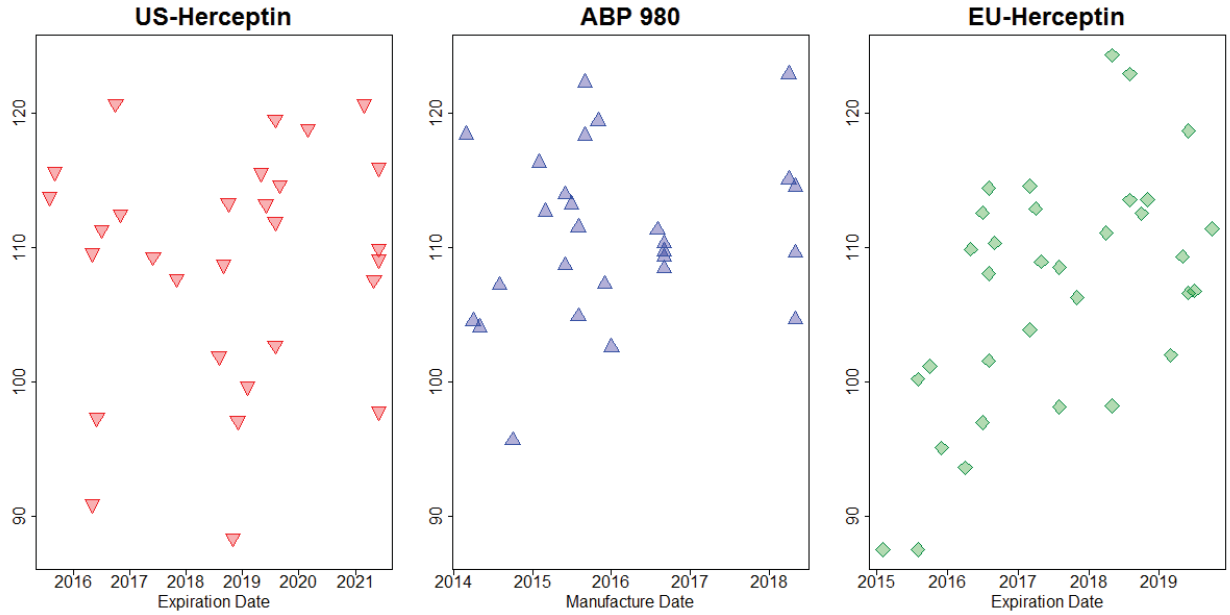
4.3 Equivalence testing for proliferation inhibition

Figure 1 shows scatter plots of proliferation inhibition for ABP 980, US-Herceptin and EU-Herceptin. Table 8 shows descriptive statistics for the inhibition of cell growth. The range of ABP 980 is similar to the range of US-Herceptin and the range of EU-Herceptin data.

Table 8 Descriptive statistics for inhibition of cell growth

Product	Number of lots	Mean, %	Sample standard deviation, %	Min, %	Max, %
ABP 980	27	111.01	6.33	95.68	122.90
US-Herceptin	29	108.83	8.60	88.42	120.71
EU-Herceptin	33	106.79	8.89	87.55	124.32

Figure 1 the proliferation inhibition bioassay



Bridging was conducted using 3-way comparisons. The three-way comparisons result for proliferation inhibition using the MWCMLE method with parameter margin are listed in Table 9. The confidence interval of mean difference falls within the EAC for each comparison. The proposed biosimilar product passes all 3-way equivalence test for the proliferation inhibition assay.

Table 9 Equivalence testing results for the proliferation inhibition using the MWCMLE for parameter margin

Test Product (Number of lots)	Comparator (Number of lots)	Mean Difference, %		Equivalence Margin, %	Conclusion
		Estimate	90% CI		
ABP 980 (27)	US (29)	2.19	(-2.91, 6.39)	(-14.61, 14.61)	Pass
ABP 980 (27)	EU (33)	4.23	(-1.59, 8.23)	(-15.11, 15.11)	Pass
EU (33)	US (29)	-2.04	(-6.58, 3.08)	(-14.61, 14.61)	Pass

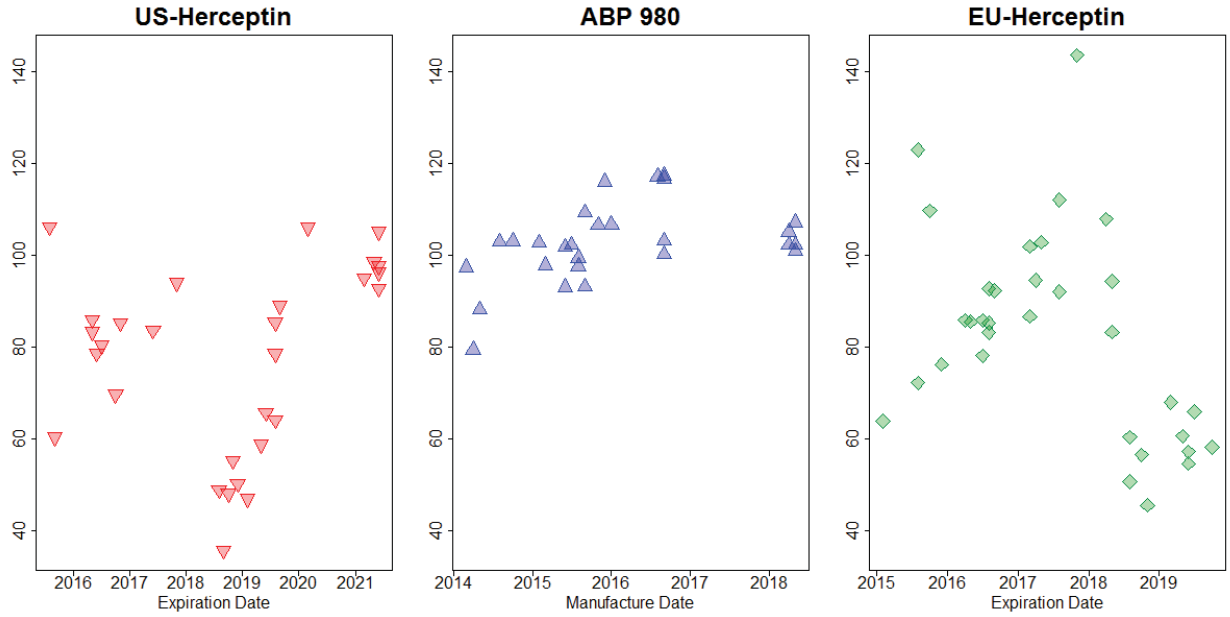
4.4 Equivalence testing for ADCC-NK92

Table 10 shows the descriptive statistics for ADCC-NK92. There are 26 US-Herceptin lots included in the analysis. The expiration dates of US-Herceptin lots range from 2015 to 2021. The expiration dates of EU-Herceptin lots range from 2015 to 2019. Figure 2 shows scatter plots of ADCC-NK92 data from ABP 980, US-Herceptin and EU-Herceptin. As shown in Figure 2, the cluster of ABP 980 is relatively higher comparing to the reference product. There is a sudden drop of ADCC-NK92 values in both US-Herceptin and EU-Herceptin.

Table 10 Descriptive statistics for ADCC-NK92

Product	Number of lots	Mean, %	Sample standard deviation, %	Min, %	Max, %
ABP 980	27	102.53	8.63	79.33	117.30
US-Herceptin	29	77.41	20.21	35.67	106.13
EU-Herceptin	33	82.70	22.40	45.53	143.47

Figure 2 ADCC-NK92



Bridging was conducted using 3-way comparisons. The three-way comparisons result for ADCC-NK92 using the MWCMLE method with parameter margin are listed in Table 11. The confidence interval of mean difference falls within the EAC for each comparison. The proposed biosimilar product passes all 3-way equivalence test for ADCC-NK92 based on the MWCMLE method.

Table 11 Equivalence testing results for ADCC-NK92 using the MWCMLE for parameter margin

Test Product (Number of lots)	Comparator (Number of lots)	Mean Difference, %		Equivalence Margin, %	Conclusion
		Estimate	90% CI		
ABP 980 (27)	US (29)	25.12	(7.70, 33.92)	(-34.36, 34.36)	Pass
ABP 980 (27)	EU (33)	19.83	(3.10, 28.27)	(-38.08, 38.08)	Pass
EU (33)	US (29)	5.29	(-7.17, 16.27)	(-34.36, 34.36)	Pass

4.5 Equivalence testing for ADCC-PBMC

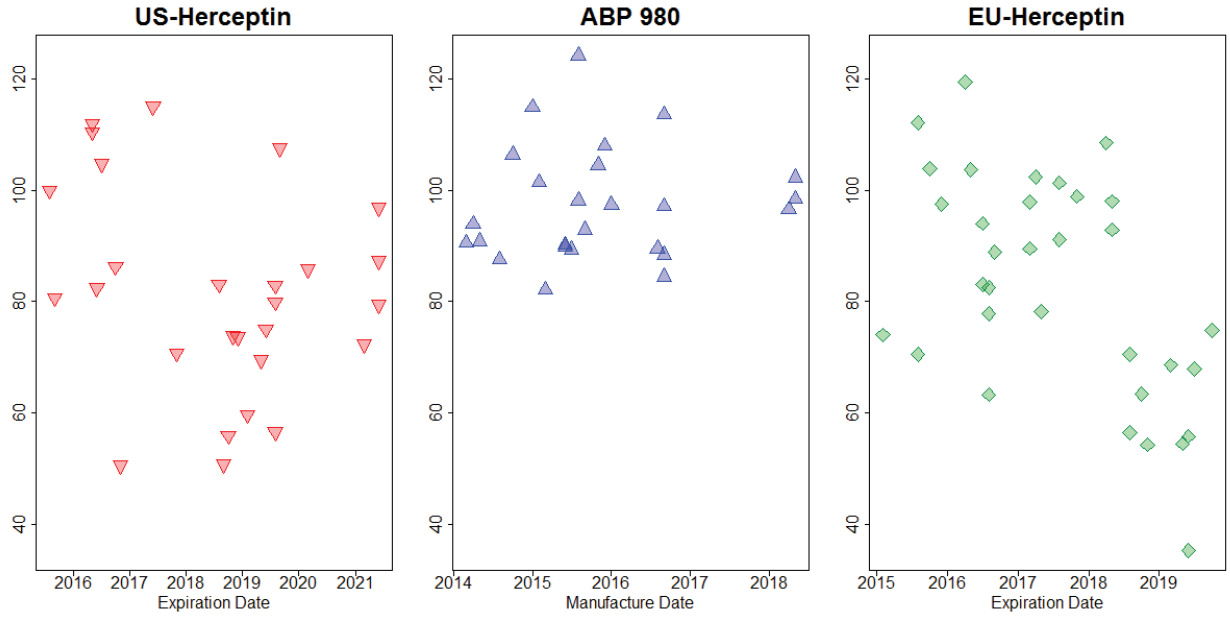
Table 12 shows the descriptive statistics for ADCC-PBMC. Figure 3 shows scatter plots of ADCC-PBMC from the three products with the same range of vertical axis. There are 27 US-Herceptin lots included in the analysis. The expiration dates of US-Herceptin lots range from 2015 to 2021. The expiration dates of EU-Herceptin lots range from 2015 to 2019. Figure 2 shows scatter plots of ADCC-NK92 data from ABP 980, US-Herceptin and EU-Herceptin. As shown in Figure 2, the cluster of ABP 980 is relatively higher comparing to the reference product. There is a sudden drop of ADCC-NK92 values in both US-Herceptin and EU-Herceptin.

In the first submission, two DP lots, 130028 (manufactured Dec 2013) and 0010163483 (manufactured Apr 2014) manufactured from the same DS lots were both test for ADCC-PBMC. The applicant suggests to use the average of values from those two lots in Tier-1 equivalence testing. The underlying distribution of the averaged value, which include an unestimable between correlated lots variation and a constant, is different from other observations. Therefore, we keep the ADCC-PBMC value from one DP lot instead of using the averaged value. Since the DP lot 130028 was manufactured earlier, we keep the value of ADCC-PBMC from DP lot 130028 in Tier-1 equivalence testing. In this resubmission, lot 130028 is no longer included in the analysis.

Table 12 Descriptive statistics for ADCC-PBMC

Product	Number of lots	Mean, %	Sample standard deviation, %	Min, %	Max, %
ABP 980	25	97.30	10.22	82.10	124.20
US-Herceptin	27	81.58	18.52	50.60	115.00
EU-Herceptin	33	82.76	19.95	35.30	119.40

Figure 3 ADCC-PBMC



Bridging was conducted using 3-way comparisons. The three-way comparisons result for ADCC-PBMC using the MWCMLE method with parameter margin are listed in Table 13. The confidence interval of mean difference falls within the EAC for each comparison. The proposed biosimilar product passes all 3-way equivalence test for the ADCC-PBMC based on the MWCMLE method.

Table 13 Equivalence testing results for ADCC-PBMC using the MWCMLE for parameter margin

Test Product (Number of lots)	Comparator (Number of lots)	Mean Difference, %		Equivalence Margin, %	Conclusion
		Estimate	90% CI		
ABP 980 (25)	US (27)	15.72	(1.41, 23.85)	(-31.48, 31.48)	Pass
ABP 980 (25)	EU (33)	14.54	(-0.18, 22.50)	(-33.92, 33.92)	Pass
EU (33)	US (27)	1.18	(-9.54, 11.58)	(-31.48, 31.48)	Pass

5 Conclusion

The results from the statistical equivalence analyses of the proliferation inhibition, ADCC-NK92 and ADCC-PBMC data support a demonstration that ABP 980 and US-Herceptin are highly similar and support the analytical portion of the scientific bridge to justify the relevance of EU-Herceptin data from the comparative clinical study.

6 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 YI HSU
05/29/2019 01:20:45 PM
link corrected

MEIYU SHEN
05/29/2019 01:22:42 PM

THOMAS J PERMUTT
05/29/2019 01:31:02 PM
I concur.

US 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

Biometrics Division: VI

BLA NO.:	761073
DATE RECEIVED BY OB:	Aug 08, 2017
DRUG NAME:	ABP 980 (proposed biosimilar to Herceptin) Kanjinti® (trastuzumab-anns)
DOSAGE FORM:	lyophilized powder in a strength of 420 mg/vial
INDICATION:	HER2 overexpressing breast cancer
APPLICANT:	Amgen
REVIEW FINISHED:	Mar 21, 2018
STATISTICAL REVIEWER:	Yu-Yi Hsu
SECONDARY REVIEWER:	Meiyu Shen
PROJECT MANAGER:	Amy Tilley

Concur: Yi Tsong, Ph.D., Division Director, CDER/OTS/OB/DB VI

Distribution:

CDER/OTS/OB/DBVI, Yi Tsong
CDER/OTS/OB/DBVI, Meiyu Shen
CDER/OTS/OB/IO, Lillian Patrician
CDER/OPQ/OBP/DBRRII, Chen Sun
CDER/OPQ/OBP/DBRRII, Patrick Lynch

Table of Contents

List of Tables

List of Figures

1 Executive summary and recommendation

The CMC statistical reviewer in the Office of Biostatistics analyzed the comparative results of two critical Quality Attributes (QAs): proliferation inhibition, and antibody-dependent cellular cytotoxicity activity using an engineered NK92 cell line (ADCC-NK92) which were recommended for equivalence testing analysis by the Office of Biotechnology Products (OBP). Tier 1 statistical equivalence testing was conducted using equivalence margins of ±1.5 times of the standard deviation of the reference product, estimated by $\hat{\sigma}_R$, where the subscript R represents the comparator product. Samples from 22 lots of ABP 980 (Amgen), 23 lots of US-licensed Herceptin, and 33 lots of Herceptin sourced at EU were used for evaluating the similarity of tier-1 QAs using the equivalence testing. Six additional US-Herceptin lots were added to ADCC-NK92 analysis in the response of the IR sent in February 08, 2018. The results are summarized in Table 1 and Table 2.

Table 1 Results of equivalence testing for proliferation inhibition

Test Product (Number of lots)	Comparator (Number of lots)	Mean Difference, %		Equivalence Margin, %	Conclusion
		Estimate	90% CI		
ABP 980 (21)	US (23)	2.31	(-1.57, 6.20)	(-13.38, 13.38)	Pass
ABP 980 (21)	EU (33)	3.99	(0.49, 7.50)	(-13.34, 13.34)	Pass
EU (33)	US (23)	-1.68	(-5.74, 2.38)	(-13.38, 13.38)	Pass

Table 2 Results of equivalence testing for the ADCC-NK92 activity

Test Product (Number of lots)	Comparator (Number of lots)	Mean Difference, %		Equivalence Margin, %	Conclusion
		Estimate	90% CI		
ABP 980 (21)	US (29)	26.47	(19.17, 33.77)	(-30.79, 30.79)	Not Pass
ABP 980 (21)	EU (33)	20.98	(13.43, 28.53)	(-33.60, 33.60)	Pass
EU (33)	US (29)	5.49	(-3.62, 14.60)	(-30.79, 30.79)	Pass

As shown in Tables 1 , the proliferation inhibition passes the 3-way equivalence test. However, the mean difference in ADCC-NK92 activity between the proposed biosimilar ABP 980 and the US-Herceptin does not pass the equivalence test as shown in Table 2.

Additional information: in the response of the IR sent in February 08, 2018, the applicant provided data for ADCC testing using a peripheral blood mononuclear cell (ADCC-PBMC) as additional information for similarity assessment. The ADCC-PBMC assay passes the 3-way equivalence test as shown in Table 3. Please see OBP's review for details about ADCC.

Table 3 Results of equivalence testing for the ADCC-PBMC assay

Test Product (Number of lots)	Comparator (Numer of lots)	Mean Difference, %		Equivalence Margin, %	Conclusion
		Estimate	90% CI		
ABP 980 (21)	US (23)	15.44	(7.25, 23.63)	(-29.58, 29.58)	Pass
ABP 980 (21)	EU (33)	13.84	(6.49, 21.2)	(-29.93, 29.93)	Pass
EU (33)	US (23)	1.60	(-7.43, 10.62)	(-29.58, 29.58)	Pass

2 Overview

ABP 980 is a proposed biosimilar to the US-licensed Herceptin (trastuzumab) submitted under Section 351(k) of the Public Health Service Act. The applicant submitted analytical similarity assessments using a three-way bridge approach to demonstrate analytical similarity between ABP 980 and the US-Herceptin and to establish a scientific bridge between the US-licensed Herceptin and EU-approved Herceptin.

Three critical quality attributes, proliferation inhibition, HER2 binding by ELISA and antibody-dependent cellular cytotoxicity activity using an engineered NK92 cell line (ADCC-NK92) were initially assigned as Tier-1 and were tested using samples from multiple biosimilar, US-Herceptin (440 mg presentation) and EU-Herceptin (150 mg presentation) lots. In the IR sent by OBP at 19th March, 2018, the HER2 binding by ELISA assay is suggested to be evaluated using Tier-2 quality range criteria in Question 1.b.

In the response of OBP's IR, "response to 08 February 2018 FDA information request – ADCC Tier 1 Analysis", the applicant tested 6 additional US-Herceptin lots for ADCC-NK92 activity. The application also provided data for ADCC activity using a peripheral blood mononuclear cell (ADCC-PBMC) in the same IR response to assist the similarity assessment.
Our comments reagarding the applicant's statistical equivalence testing for proliferation inhibition, ADCC-NK92 and ADCC-PBMC are provided in Section 3. Our independent statistical analyses and descriptions of received data are summarized in Section 4. Our conclusion is provided in Section 4.4.

3 Applicant's statistical equivalence testing

In this submission, the applicant 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 comparator product lots, for proliferation inhibition, ADCC-NK92, and ADCC-PBMC.

In addition to the FDA recommended method, the applicant also proposed another equivalence test based on medians to demonstrate the difference between the medians of the test product and the comparator product. The median based method uses quantile regression to calculate 90% confidence interval and uses ±1.5*inter quantile range as EAC.

The applicant observed a shift in ADCC-NK92 activity for US-Herceptin and EU-Herceptin lots with expiration date beyond July 2018. The applicant defined US-Herceptin and EU-Herceptin with expiration date beyond July 2018 as post-shift lots. The applicant claim that the shift caused lower sample mean of ADCC-NK92 for US-Herceptin. For that reason, the applicant applied both the mean based and the median based methods for ADCC-NK92 and claim that the median based method is less affected by the shift. However, it is not clear if using the multiplier 1.5 combing with IQR as the equivalence margin is justifiable. Furthermore, comparison in medians provides different interpretation from the comparison of means in terms of central tendency.

The CMC statistical reviewer performs an independent statistical analysis in the next section. We use all available data that includes data from all available lots for tests of Tier-1 QAs: proliferation inhibition, HER2 binding by ELISA, and ADCC-NK92. We also perform equivalence test using all available data for the additional quality attribute ADCC-PBMC.

4 Analytical similarity

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. More details are described in the FDA Draft Guidance on Analytical Similarity (2017)[1]. This review focuses on the Tier 1 statistical equivalence testing

4.1 Data analyzed

The applicant tested multiple lots of ABP 980, US- Herceptin and EU- Herceptin for Tier-1 QAs. The numbers of lots tested for Tier-1 QAs are listed in Table 4.

Table 4 Number of lots tested for Tier-1 quality attributes

Quality attribute	ABP 980	US-Herceptin	EU-Herceptin
Proliferation inhibition	21	23	33
ADCC activity using NK92	21	29	33

The applicant provided data from 22 biosimilar lots, including 7 drug product lots with 150mg presentation (DP-150mg), 6 drug product lots with 440mg presentation (DP-440mg) and 9 drug substance (DS) lots for assessing analytical similarity. According to Table 1 of "Analytical similarity assessment – testing plan and assessment criteria", two DP-150mg Lots, ABP 130030 and ABP 0010204750, were manufactured using the same drug substance lot. Since ABP 130030 has earlier manufacturing date, we keep ABP-130030 in the analysis data set. The remaining 21 ABP lots are corresponding to 21 different drug substance numbers.

Table 5 shows descriptive statistics for the proliferation inhibition bioassay data grouped by lot type (DS or DP).

Figure 1 shows the proliferation inhibition data against manufacture date. One EU lot, EU-H4499H02, has relative potency rounded to the nearest one on Table 1 of Section "analytical similarity assessment – biological activity". The other relative potency values are rounded to the third digits on Table 1 of Section "recorded datasets used for statistical analysis".

Table 5 Descriptive statistics for the proliferation inhibition bioassay of ABP 980 grouped by type (DS or DP)

Subgroup	Number of lots	Mean, %	Sample standard deviation, %	Min, %	Max, %
DP – 150mg	6	109.84	8.16	95.68	118.43
DP – 440mg	6	107.50	3.16	102.61	110.34
DS	9	113.60	5.67	104.92	122.31
All (DP+DS)	21	110.78	6.25	95.68	122.31

Figure 1 Proliferation inhibition bioassay of ABP 980

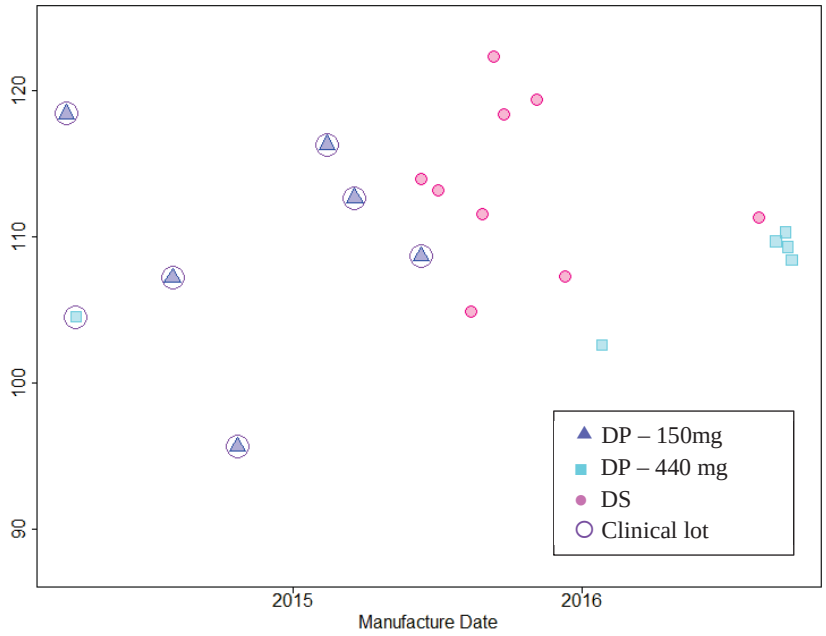


Table 6 shows descriptive statistics for ADCC using NK cell for ABP 980 lots grouped by lot types. The patterns are similar between groups. Figure 2 shows the ADCC-NK92 against manufacturing dates. All 21 lots tested for inhibition are also tested for ADCC-NK92.

Figure 2 ADCC-NK92 for ABP 980

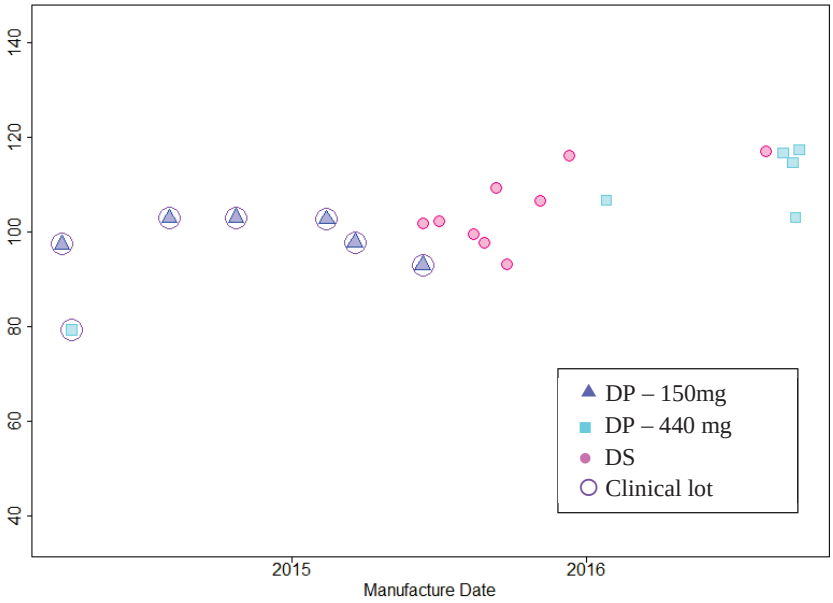


Table 6 Descriptive statistics for the ADCC-NK92 of ABP 980 grouped by type (DS or DP)

Subgroup	Number of lots	Mean, %	Sample standard deviation, %	Min, %	Max, %
DP – 150mg	6	99.46	4.10	92.97	103.00
DP – 440mg	6	106.26	14.38	79.33	117.30
DS	9	104.77	8.17	93.07	117.10
All (DP+DS)	21	103.67	9.51	79.33	117.30

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 $\sigma_R$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 test the following null and alternative hypotheses:

$$H_0: \mu_T - \mu_R \leq \theta_1 \text{ or } \mu_T - \mu_R \geq \theta_2$$
$$H_1: \theta_1 < \mu_T - \mu_R < \theta_2$$

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

The null hypothesis is rejected with significance level not exceeding α=0.05 if the 90% Confidence Interval (CI) of the mean difference, $\mu_T - \mu_R$, of the QA of interest falls within $(-1.5\sigma_R, 1.5\sigma_R)$. In other words, we conclude equivalence of the QA of interest between the test product and the reference product if null hypothesis is rejected. Since the margin is conducted with unknown parameter, we replace σ_R by the sample standard deviation of the reference product.

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.

Let $\bar{X}_i = \sum_{j=1}^{n_i} X_{Tj}/n_i$ and $S_i^2 = \sum_{j=1}^{n_i}\left(X_{ij} - \bar{X}_i\right)^2/(n_i - 1)$, where n_i is the number of lots of the product i, and $i \in \{T, R\}$. Assuming unequal variances of the test product and the reference product, the $(1-2\alpha) * 100\%$ CI of the mean difference in the QA of interest can be calculated as $\left(\bar{X}_T - \bar{X}_R - t_\alpha(\nu)\sqrt{S_T^2/n_T^* + S_R^2/n_R^*}, \bar{X}_T - \bar{X}_R + t_\alpha(\nu)\sqrt{S_T^2/n_T^* + S_R^2/n_R^*}\right)$, where $t_\alpha(\nu)$ is the $1-\alpha$ t-distribution quantile and $\nu = \frac{\left(S_T^2/n_T^* + S_R^2/n_R^*\right)^2}{\frac{(S_T^2/n_T^*)}{(n_T-1)} + \frac{(S_R^2/n_R^*)}{(n_R-1)}}$ is the degrees of freedom calculated by Welch-Satterthwaite's approximation with adjustment for imbalance sample size. The adjusted value $n_R^* = 1.5n_T$ if $n_R > 1.5n_T$ and $n_R^* = n_R$ otherwise, and $n_T^* = 1.5n_R$ if $n_T > 1.5n_R$ and $n_T^* = n_T$ otherwise.

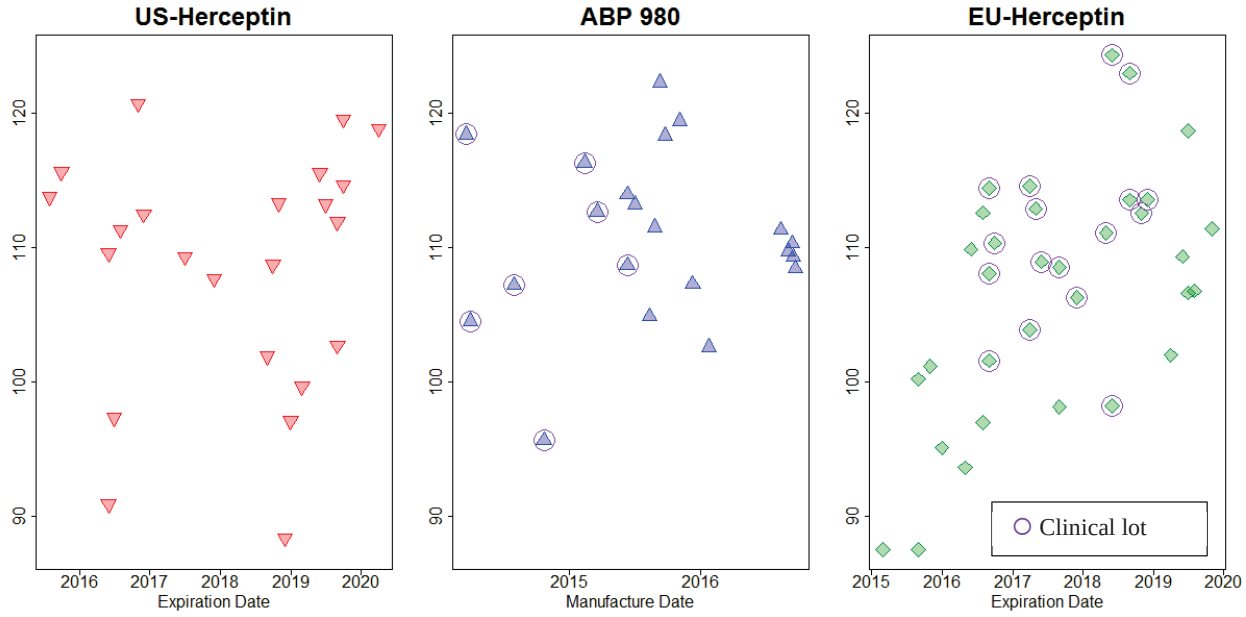
4.3 Equivalence testing for proliferation inhibition

Figure 3 shows scatter plots of proliferation inhibition for ABP 980, US-Herceptin and EU-Herceptin. Table 7 shows descriptive statistics for the inhibition of cell growth. The range of ABP 980 is similar to the range of US-Herceptin and the range of EU-Herceptin data.

Table 7 Descriptive statistics for inhibition of cell growth

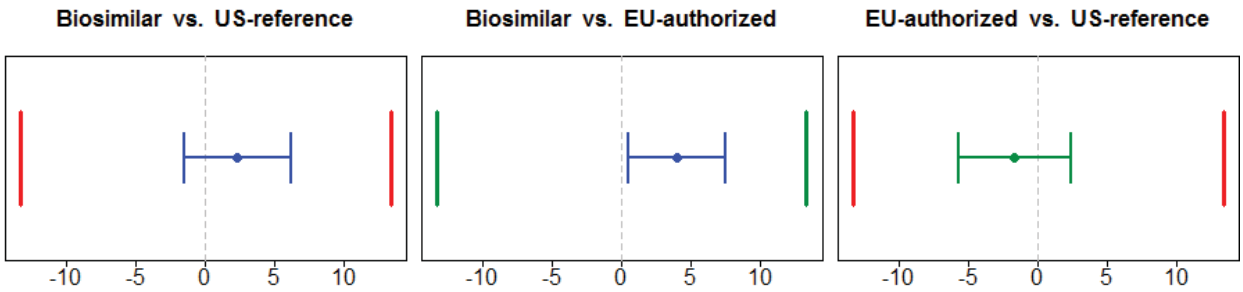
Product	Number of lots	Mean, %	Sample standard deviation, %	Min, %	Max, %
ABP 980	21	110.78	6.25	95.68	122.31
US-Herceptin	23	108.47	8.92	88.42	120.71
EU-Herceptin	33	106.79	8.89	87.55	124.32

Figure 3 the proliferation inhibition bioassay



Bridging was conducted using 3-way comparisons. The 3-way equivalence testing results are summarized in Table 1. The graphical expressions of the 3-way equivalence testing are shown in Figure 4. As shown in Figure 4, the confidence interval of mean difference falls within the EAC for each comparison. The proposed biosimilar product pass all 3-way equivalence test for the proliferation inhibition assay.

Figure 4 Equivalence margin and 90%CI of the proliferation inhibition bioassay



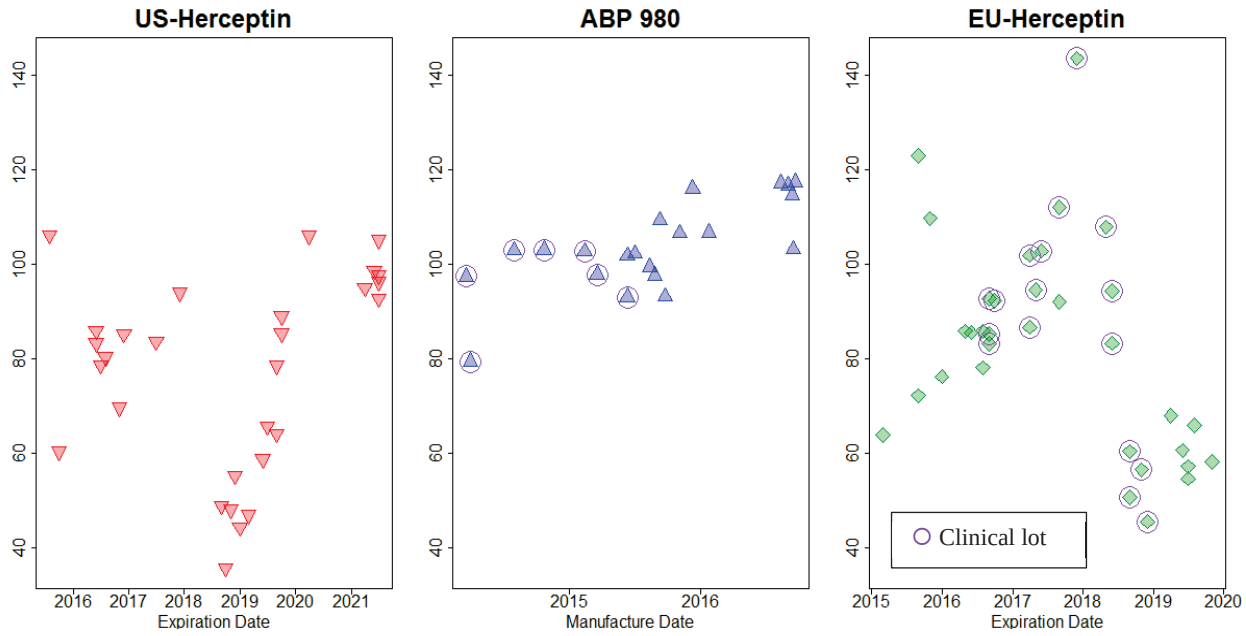
4.4 Equivalence testing for ADCC-NK92

shows the descriptive statistics for ADCC-NK92. There are 29 US-Herceptin lots included in the analysis, including 6 additional lots provided in the IR response "response to 08 February 2018 FDA information request – ADCC Tier 1 Analysis". The expiration dates of US-Herceptin lots range from July 2015 to June 2021. The expiration dates of EU-Herceptin lots range from February 2015 to October 2019. Figure 5 shows scatter plots of ADCC-NK92 data from ABP 980, US-Herceptin and EU-Herceptin. As shown in Figure 5, the cluster of ABP 980 is relatively higher. There is a sudden drop of ADCC-NK92 values in both US-Herceptin and EU-Herceptin.

Table 8 Descriptive statistics for ADCC-NK92

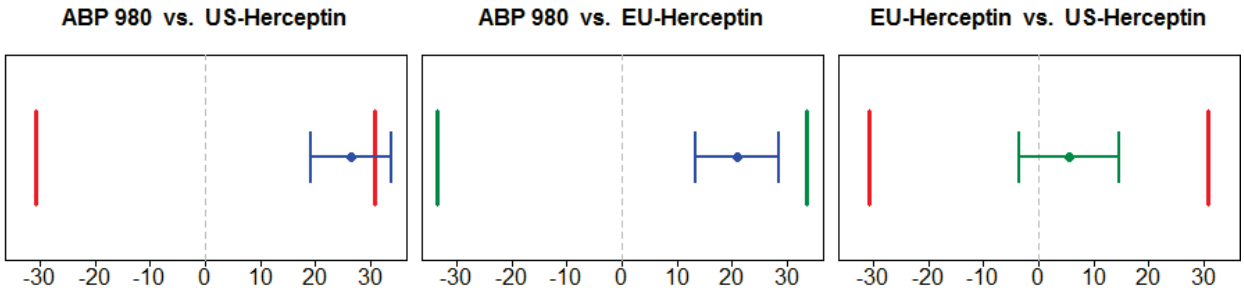
Product	Number of lots	Mean, %	Sample standard deviation, %	Min, %	Max, %
ABP 980	21	103.67	9.51	79.33	117.30
US-Herceptin	29	77.20	20.53	35.67	106.13
EU-Herceptin	33	82.70	22.40	45.53	143.47

Figure 5 ADCC-NK92



The 3-way equivalence testing results are listed in Table 2. The test results are also graphically displayed in Figure 6. As shown in the first panel of Figure 6, the proposed biosimilar does not pass equivalence test for ADCC-NK92 assay. Which means that we don't have enough evidence to reject the null hypothesis that the proposed biosimilar product and the US reference product may have ADCC-NK92 mean difference above the EAC upper bound.

Figure 6 Equivalence margin and 90%CI of ADCC-NK92 assay



4.5 Additional information: equivalence testing for ADCC-PBMC

The applicant provided additional information for ADCC-PBMC from 21 ABP980 lots, 23 US lots and 28 EU lots in the response of OBP's IR, "response to 08 February 2018 FDA information request – ADCC Tier 1 Analysis". In the later response of OBP's IR, "Response to Questions Dated 19 March 2018 – Analytical similarity, control strategy, reference standard and expiry", information of ADCC-PBMC from 6 EU lots were provided.

Two DP lots, 130028 (manufactured Dec 2013) and 0010163483 (manufactured Apr 2014) manufactured from the same DS lots were both test for ADCC-PBMC. The applicant suggests to

use the average of values from those two lots in Tier-1 equivalence testing. The underlying distribution of the averaged value, which include an unestimable between correlated lots variation and a constant, is different from other observations. Therefore, we keep the ADCC-PBMC value from one DP lot instead of using the averaged value. Since the DP lot 130028 was manufactured earlier, we keep the value of ADCC-PBMC from DP lot 130028 in Tier-1 equivalence testing.

There were 23 US-Herceptin lots tested for ADCC-PBMC. The additional 6 lots tested for ADCC-NK92 are not tested for ADCC-PBMC at the time of response. The set of 33 EU-Herceptin lots tested for ADCC-PBMC were also tested for ADCC-NK92. Table 9 shows the descriptive statistics for ADCC-PBMC. **Error! Not a valid bookmark self-reference.** shows scatter plots of ADCC-PBMC from the three products with the same range of vertical axis. Figure 8 shows the relationship between ADCC-NK92 and ADCC-PBMC.

Table 9 Descriptive statistics for ADCC-PBMC

Product	Number of lots	Mean, %	Sample standard deviation, %	Min, %	Max, %
ABP 980	21	96.60	11.81	78.00	124.20
US-Herceptin	23	81.16	19.72	50.60	115.00
EU-Herceptin	33	82.76	19.95	35.30	119.40

Figure 7 ADCC-PBMC

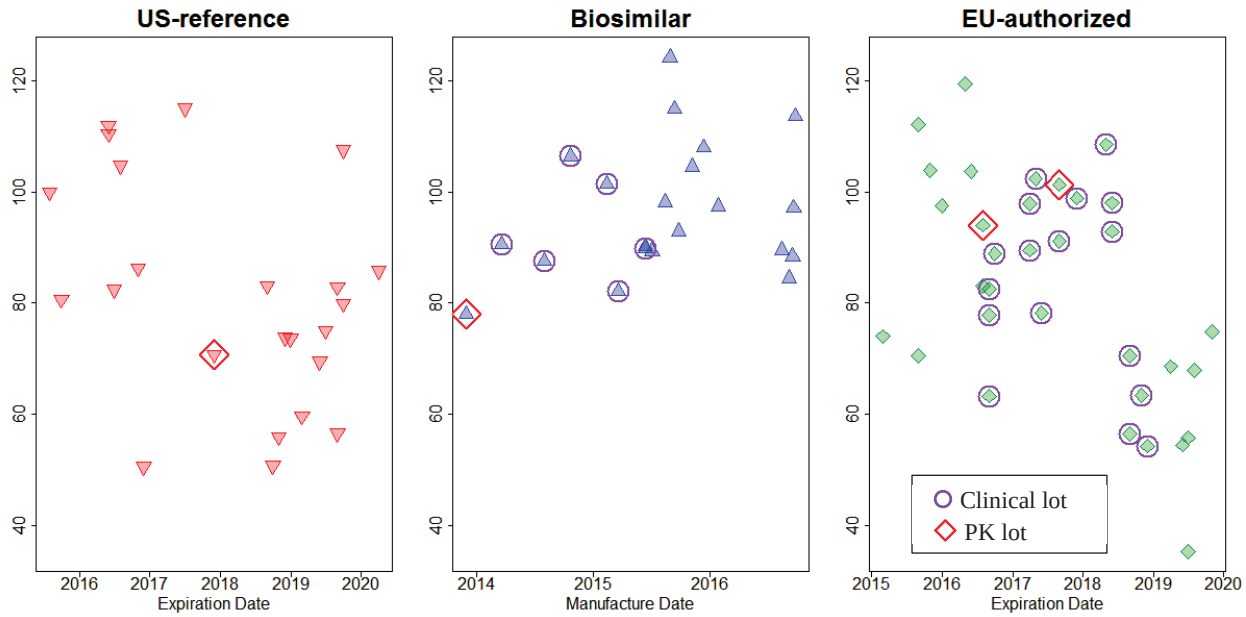
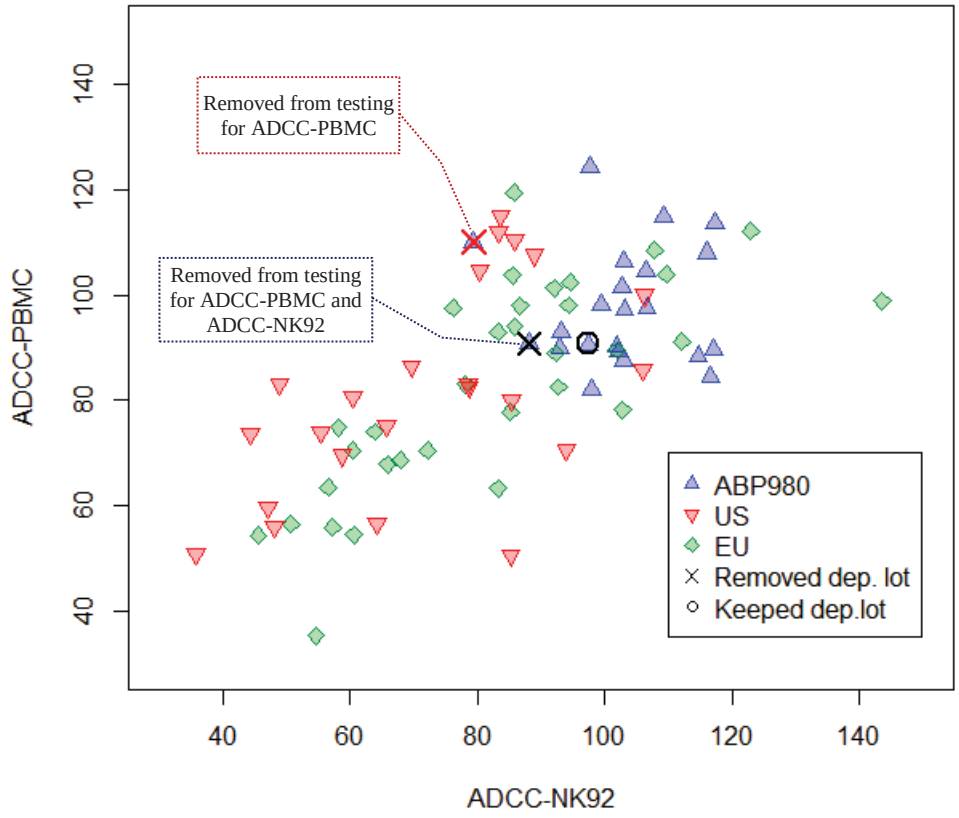
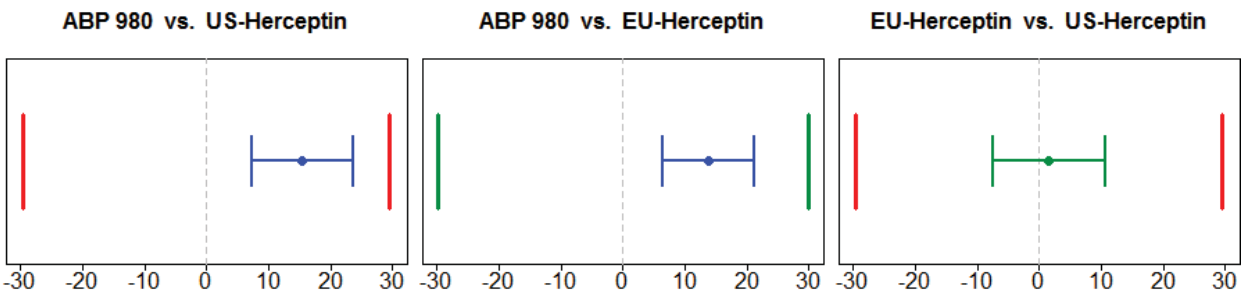


Figure 8 Plot of ADCC-NK92 and ADCC-PBMC



We conducted the equivalence test on ADCC-PBMC. The 3-way equivalence testing results are listed in Table 3. The test results are also graphically displayed in Figure 9. As shown in Figure 9, the confidence interval of mean difference falls within the EAC for each comparison. The proposed biosimilar product pass all 3-way equivalence test for the ADCC-PBMC assay.

Figure 9 Equivalence margin and 90%CI of ADCC-PBMC assay



5 Conclusion

The results from the statistical equivalence analyses of the proliferation inhibition and ADCC-PBMC data support a demonstration that ABP 980 and US-Herceptin are highly similar and support the analytical portion of the scientific bridge to justify the relevance of EU-Herceptin data from the comparative clinical study. However, ADCC-NK92 activity does not pass the equivalence test for comparing ABP 980 and US-Herceptin. The mean difference in ADCC-NK92 between ABP 980 and US-Herceptin is 26.47%.

6 Reference

[1] Scientific considerations in demonstrating biosimilarity to a reference product (2015) https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf

[2] 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 on behalf of YU YI HSU
04/24/2018

MEIYU SHEN
04/24/2018

YI TSONG
04/24/2018