CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

125409Orig1s000

STATISTICAL REVIEW(S)

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

STATISTICAL REVIEW AND EVALUATION

BLA Number:	125409
Drug Name:	Pertuzumab
Indication(s):	Metastatic Breast Cancer
Applicant:	Genentech, Inc.
Date(s):	Date Received: 1/27/2012 Completion Date: 6/5/2012
Biometrics Division:	VI
Statistical Reviewer:	Youngsook Jeon, Ph.D. (OTS/OB/DBVI)
Concurring Reviewer:	Yi Tsong, Ph.D. (OTS/OB/DBVI)
Consult Requesting Reviewer:	Laurie Graham, Ph.D. (OPS/OBP/DMA)
Project Manager:	Amy Tilley (OND/OHOP/DOP1)

Distribution: BLA 125409
OTS/OB/ Y. Tsong, S. Machado, L. Patrician
OPS/OBP/DMA/L. Graham
OND/OHOP/DOP1/A. Tilley

Table of Contents

1. EXCECUTIVE SUMMARY

This review describes statistical findings on Genentech's equivalence test between small-scale and manufacturing-scale models for Pertuzumab so that FDA Office of Biotechnology Product (OBP) can make an informed decision.

Genentech used scale-down models to predict performance at manufacturing scale. In order to evaluate the equivalency between small-scale and manufacturing-scale models, Genentech conducted equivalence tests for 1) (b) (4) Regarding to Genentech's equivalence tests, Dr. Graham in OBP requested the statistical consulting.

In response to the consulting request, FDA statistical reviewers reviewed Genentech's submission. Reviewers found the following three issues. First, there is not enough data for (b) (4) To gain enough statistical power, more data are necessary for an equivalence test. Second, it is not acceptable that (b) (4) . Last, for both Probable Non Equivalence and Not Equivalent attributes, an offset was not validated by independent data.

2. INTRODUCTION

For process development and full process characterization, Genentech used scale-down models, which enable to predict performance at manufacturing scale. In order to evaluate the equivalency between small-scale and manufacturing-scale models, Genentech conducted a statistical test for equivalence for (b) (4)

Genentech conducted equivalence tests against a practically significant difference (PSD). The PSD is an interval over which a difference between the test groups was not practically significant. For KPIs and non-CQAs, the PSD is (b) (4) of the manufacturing scale average. For CQAs with a defined CQA-TR, PSD is (b) (4) of the difference between the manufacturing scale average and the CQA-TR.

Regarding to this statistical content, Dr. Laurie Graham in DMA/OBP/OPS/CDER/FDA requested the following statistical consulting:

Use of data from small-scale models of biotechnology product manufacturing processes is essential for process development and full process characterization. Critical to the review of data from small-scale models is the need to determine how well the models represent the commercial manufacturing process. The sponsor has provided in the submission a statistical evaluation of equivalency between the small-scale models and commercial manufacturing unit operations. We would like the statistical analysis of the sponsor evaluated to determine if it will provide adequate assurance that the output of the small-scale models are representative of the commercial manufacturing process.

In response to the above consulting request, FDA statisticians reviewed Genentech's submission with focus on their equivalence evaluation between small-scale and manufacturing-scale models under the section: Qualifications of the cell culture and purification scale-down models in 3.2.S.2.5 Process Validation and/or Evaluation (pages 35-63). Section 3 describes the reviewer's assessment. In Section 3, the reviewers redefine Genentech's categories for equivalence testing outcomes because it is inconsistent with those for traditional equivalence testing practice. Section 4 discusses summary and conclusions.

3. REVIEWER'S ASSESSMENT

3.1 Sponsor's Data

Genentech conducted equivalence tests for 1) (b) (4) [redacted] There is not enough statistical power to establish the equivalence.

Table 1 shows the summary statistics of the data used for Cell Culture Scale Qualification study. Other sponsor's data sets and summary statistics can be found in Tables 7 – 10 in the sponsor's report, 3.2.S.2.5 Process Validation and/or Evaluation (pages 24 – 35).

Table 3.2.S.2.5-7 (b) (4)

Table 1: Table 7 in Sponsor's Report, 3.2.S.2.5 Process Validation and/or Evaluation (page 35 of 315)

3.2 Statistical Evaluation

Sponsor's Statistical Approach: Equivalence Test

Genentech conducted equivalence test by calculating 90% confidence intervals for the mean difference between manufacturing-scale and small-scale data assuming unequal variances against a practically significant difference (PSD). PSD is defined as an interval over which a difference between the test groups was not considered different. PSD is generally taken to be 10 % of manufacturing scale average for KPI[1]s and non-CQA[2]s and 10% of the allowed variation for CQAs with some exceptions. Detailed explanation for PSD is summarized in another sponsor's report, VP08-137 (b) (4) and is not reviewed here.

According to the equivalence testing results, the attributes were classified into for categories by the sponsor: Equivalent, Probable Equivalence, Probable Non-Equivalence, and Not Equivalent. An attribute is Equivalent if 90% confidence interval is entirely within PSD. If it is not Equivalent

[1] Key Performance Indicator
[2] Critical Quality Attribute

but the mean difference is within PSD, the attribute is Probable Equivalence. If it is not Equivalent and the mean difference is outside PSD, the attribute is Probable Non-Equivalence. If 90% confidence interval is entirely outside PSD, the corresponding attribute is Not Equivalent.

In sum, the sponsor made the following conclusions based on the equivalence test with some exceptions:

1. If Equivalent, the small-scale model was qualified without further evaluation.
2. If Probable Equivalence, the small-scale model was qualified because 1) 90% confidence intervals are narrower than PSD ranges, 2) the variations in different scales are similar, and 3) the mean difference is within PSD.
3. If Probable Non-Equivalence or Not Equivalent, the small-scale model was qualified with an offset, which equals to the observed difference in means, because 1) 90% confidence intervals are narrower than PSD ranges; 2) the variations in different scales are similar; and 3) the observed mean difference is smaller than variations observed in multivariate experiment results in small-scale models.

Although this categorization was agreed by the previous FDA statistical reviewer (Dr. Zhong, Division of Biometrics I, Office of Biostatistics), it is inconsistent with the definition used in usual equivalence testing practice including bioequivalence. Therefore, we redefine the equivalency categorization this time. An attribute is Equivalent if 90% confidence interval is entirely within PSD. If it is not Equivalent but the mean difference is within PSD, the attribute is called Equivalent-in-Sample-Mean-Only. If it is not Equivalent and the mean difference is outside PSD, the attribute is Failed-to-be-Equivalent. If 90% confidence interval is entirely outside PSD, the corresponding attribute is Inequivalent. We will use the new defined categories here.

Two Issues in Sponsor's Conclusion

The reviewer found two issues in the sponsor's general conclusions.

First, there is no difference in the sponsor's conclusions regarding model qualification between Equivalent and Equivalent-in-Sample-Mean-Only attributes. The equivalence is not shown for the Equivalent-in-Sample-Mean-Only attribute because the 90% confidence interval is not entirely within the equivalence limit (i.e. PSD). Nevertheless, the sponsor claimed that the small-scale model was qualified without an offset for the Probable Equivalence attribute just as for the Equivalent attribute. Neither further investigation nor evaluation was taken. The sponsor rationalized that the 90% confidence intervals are narrower than PSD ranges and the mean difference is within PSD limits.

However, Equivalent-in-Sample-Mean-Only attributes are not same as Equivalent attributes and need to be treated differently. Treating Equivalence-in-Sample-Mean-Only attribute same as Equivalent attribute may cause serious bias in the use of scale-down model for predicting performance at manufacturing scale in the future. This bias could be more problematic for some Equivalence-in-Sample-Mean-Only attributes whose 90% confidence intervals are mostly outside the PSD. Such attributes are ivPCV and G0 in Figure 1 - Figure 2 and CHOP in Figure 3 in APPENDIX.

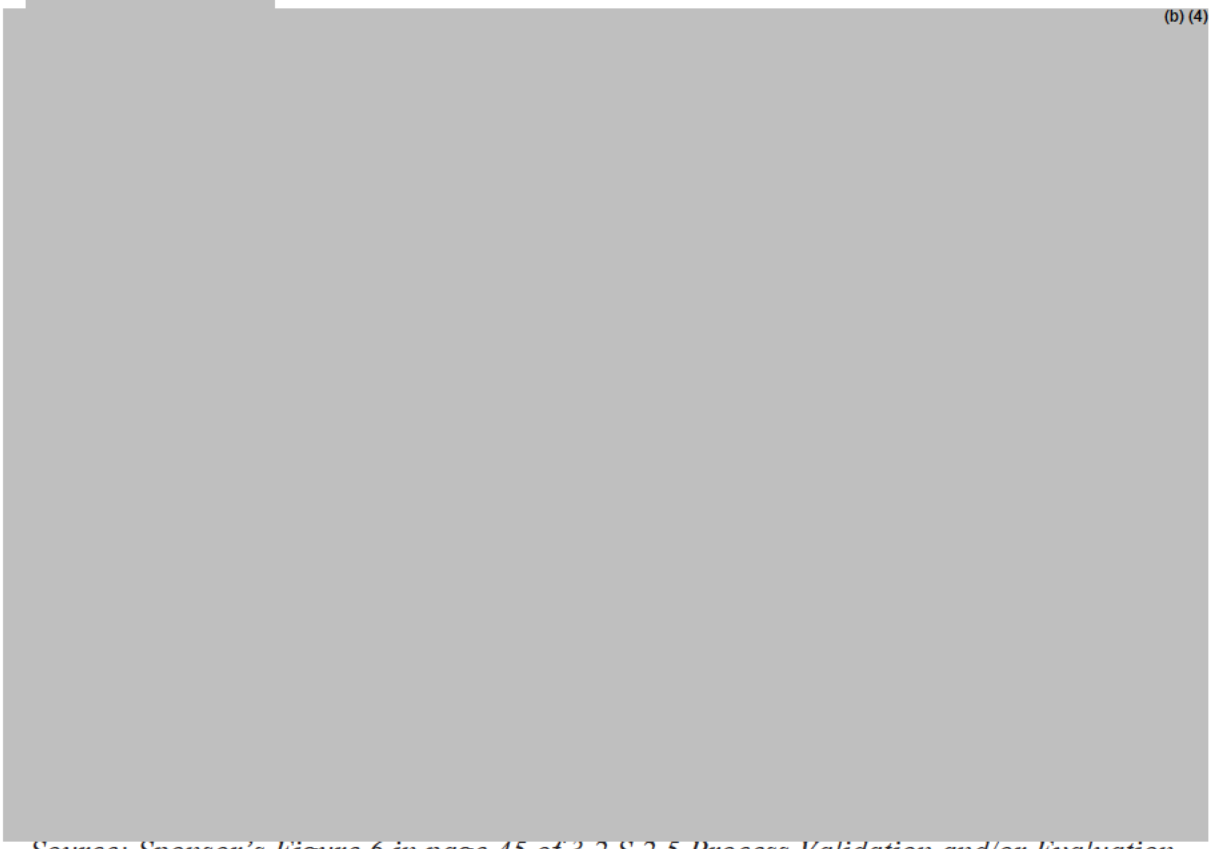
The second issue is in the sponsor's conclusion regarding Failed-to-be-equivalent and Inequivalent attributes. For such attributes, model offsets were defined and the scale-down models were qualified with offset. Model offsets are the observed mean differences. In her earlier comments to OBP (8/5/2011), Dr. Zhong pointed out that this offset, which is used to eliminate the observed bias, needs to be validated by independent data. However, the model-offset validation was not included in Genentech's report. An analysis using independent data must be conducted to ensure that an applied offset truly eliminates the bias.

4. SUMMARY AND CONCLUSIONS

FDA statisticians reviewed Genentech's submission with focus on the equivalence test between small-scale and manufacturing-scale models. Reviewers found three issues. Firstly, there is not enough data for (b) (4) study (b) (4) . To gain enough statistical power, more data are necessary for the equivalence test. Secondly, it is not acceptable (b) (4) Thirdly, for both Failed-to-be-Equivalent and Inequivalent attributes, model offsets were not validated by independent data. An analysis using independent data must be conducted to ensure that an applied offset truly eliminates the bias.

5. APPENDIX

Figure 1. (b) (4)



Source: Sponsor's Figure 6 in page 45 of 3.2.S.2.5 Process Validation and/or Evaluation

Figure 2. (b) (4)

Source: Sponsor's Figure 6 in page 45 of 3.2.S.2.5 Process Validation and/or Evaluation

Figure 3. (b) (4)

Source: Sponsor's Figure 9 in page 45 of 3.2.S.2.5 Process Validation and/or Evaluation

---

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

/s/

YOUNGSOOK JEON
06/06/2012

YI TSONG
06/06/2012

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-TEAM LEADER'S MEMO

NDA/BLA Serial Number:	125-409 / S-000
Drug Name:	Perjeta® (pertuzumab)
Indication(s):	Treatment of patients with first line HER-2 positive metastatic breast cancer.
Applicant:	Genentech, Inc.
Date(s):	Submitted: December 6, 2011 PDUFA: June 8, 2012 Review Completed: May 10, 2012
Review Priority:	Priority
Biometrics Division:	Division of Biometrics V (HFD-711)
Primary Reviewer:	Somesh Chattopadhyay, Ph.D.
Secondary Reviewer:	Shenghui Tang, Ph.D., Team Leader
Concurring Reviewer:	Thomas Gwise , Ph.D., Deputy Director
Medical Division:	Division of Oncology Products 1 (HFD-150)
Clinical Team:	Gideon Blumenthal, M.D., Medical Reviewer Nancy Scher, M.D., Medical Reviewer Patricia Cortazar, M.D., Medical Team Leader
Project Manager:	Ms. Amy Tilley

Keywords: Intent-to-treat, interim analysis, Kaplan-Meier product limit, logrank test, multiple endpoints, proportional hazards, randomization, stratification, subgroup analysis, survival analysis.

The applicant has submitted results from one multicenter, phase III, randomized, double-blind clinical trial (Study WO20698/TOC4129g or CLEOPETRA) comparing pertuzumab, a new molecular entity (NME) in combination with trastuzumab and docetaxel, to placebo in combination with trastuzumab and docetaxel in patients with HER2-positive locally recurrent, unresectable or metastatic breast cancer (MBC) who have not received anti-cancer treatment for their metastatic disease (except a maximum of one prior hormonal treatment for MBC).

In the CLEOPETRA study patients were randomized in a 1:1 ratio to receive either pertuzumab intravenously at a loading dose of 840 mg/kg followed by a dose of 420 mg/kg every 3 weeks or pertuzumab placebo intravenous infusion every 3 weeks. All patients received intravenous trastuzumab at a loading dose of 8 mg/kg followed by 6 mg/kg every 3 weeks and intravenous docetaxel at a dose of 75 mg/m2 every 3 weeks for at least 6 cycles. The randomization was stratified by prior treatment status (de novo vs. adjuvant or neoadjuvant) and region (Europe vs. North America vs. South America vs. Asia). The study started on February 12, 2008. The data cut-off date was May 13, 2011. A total of 808 patients were randomized, 402 to the experimental arm and 406 to the control arm. Patients were enrolled at 204 centers in 25 countries. The primary efficacy endpoint was progression-free survival (PFS) as assessed by an independent radiology facility (IRF). The secondary efficacy endpoints included overall survival (OS), investigator-assessed PFS, overall response rate (ORR) as assessed by the IRF, and duration of response. The pertuzumab+trastuzumab+docetaxel arm showed statistically significant improvement over placebo+trastuzumab+docetaxel arm with respect to IRF-assessed PFS in the intent-to-treat (ITT) population [hazard ratio=0.618, 95% confidence interval: (0.510, 0.749), log-rank test stratified by prior treatment status and region, two-sided p-value<0.0001]. The median PFS and its 95% confidence interval in pertuzumab+trastuzumab+docetaxel and placebo+trastuzumab+docetaxel arms were 18.5 months [95% CI: (14.6 months, 22.1 months)] and 12.4 months [95% CI: (10.4 months, 13.2 months)], respectively. With this application, the applicant submitted the results of an interim OS analysis based on data cut-off of 13 May, 2011. The results of the interim OS analysis in the ITT population showed that there was no statistically significant difference between the two treatment arms with respect to OS (log-rank test stratified by prior treatment status and region, nominal two-sided p-value 0.0053). The p-value did not cross the O'Brien-Fleming boundary which is 0.0012 with the observed number of 165 deaths (42.86% of total number of deaths required for the final OS analysis). OS medians were not reached in both arms. For further details regarding the design, data analyses, and results of the CLEOPETRA study, please refer to the statistical review by Dr. Somesh Chattopadhyay (May 10, 2012).

This team leader concurs with the recommendations and conclusions of the statistical reviewer (Dr. Somesh Chattopadhyay) of this application. The statistical results provide adequate evidence to support the PFS claim proposed in the BLA.

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

/s/

SHENGHUI TANG
05/10/2012

THOMAS E GWISE
05/10/2012

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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA Serial Number:	125-409 / S-000
Drug Name:	Perjeta® (pertuzumab)
Indication(s):	Treatment of patients with first line HER-2 positive metastatic breast cancer.
Applicant:	Genentech, Inc.
Date(s):	Submitted: December 6, 2011
	PDUFA: June 8, 2012
	Review Completed: May 10, 2012
Review Priority:	Priority
Biometrics Division:	Division of Biometrics V (HFD-711)
Statistical Reviewer:	Somesh Chattopadhyay, Ph.D.
Concurring Reviewers:	Shenghui Tang, Ph.D., Team Leader
	Thomas Gwise , Ph.D., Deputy Director
Medical Division:	Division of Oncology Products 1 (HFD-150)
Clinical Team:	Gideon Blumenthal, M.D., Medical Reviewer
	Nancy Scher, M.D., Medical Reviewer
	Patricia Cortazar, M.D., Medical Team Leader
Project Manager:	Ms. Amy Tilley

Keywords: Intent-to-treat, interim analysis, Kaplan-Meier product limit, logrank test, multiple endpoints, proportional hazards, randomization, stratification, subgroup analysis, survival analysis.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1. EXECUTIVE SUMMARY

The applicant has submitted results from one multicenter, phase III, randomized, double-blind clinical trial (Study WO20698/TOC4129g or CLEOPETRA) comparing pertuzumab, a new molecular entity (NME) in combination with trastuzumab and docetaxel, to placebo in combination with trastuzumab and docetaxel in patients with HER2-positive locally recurrent, unresectable or metastatic breast cancer (MBC) who have not received anti-cancer treatment for their metastatic disease (except a maximum of one prior hormonal treatment for MBC). The pertuzumab arm showed statistically significant improvement over the placebo arm in progression-free survival (PFS) as assessed by an independent radiology facility in all randomized patients. At the time of the analysis of PFS, the overall survival (OS) data were not mature and the pertuzumab arm did not show statistically significant improvement with respect to overall survival (OS) in an interim analysis. The statistical results provide adequate evidence to support the PFS claim proposed in the BLA.

This application is based on one Phase III trial WO20698/TOC4129g (CLEOPETRA), two supporting Phase II studies (WO20697 [NEOSPHERE] and BO17929), as well as a number of Phase I and II studies. This review is primarily based on the Phase III study. In the CLEOPETRA study patients were randomized in a 1:1 ratio to receive either pertuzumab intravenously at a loading dose of 840 mg/kg followed by a dose of 420 mg/kg every 3 weeks or pertuzumab placebo intravenous infusion every 3 weeks. All patients received intravenous trastuzumab at a loading dose of 8 mg/kg followed by 6 mg/kg every 3 weeks and intravenous docetaxel at a dose of 75 mg/m2 every 3 weeks for at least 6 cycles. The randomization was stratified by prior treatment status (de novo vs. adjuvant or neoadjuvant) and region (Europe vs. North America vs. South America vs. Asia). The study started on February 12, 2008. The data cut-off date was May 13, 2011. A total of 808 patients were randomized, 402 to the experimental arm and 406 to the control arm. Patients were enrolled at 204 centers in 25 countries. The primary efficacy endpoint was progression-free survival (PFS) as assessed by an independent radiology facility (IRF). The secondary efficacy endpoints included overall survival (OS), investigator-assessed PFS, overall response rate (ORR) as assessed by the IRF and duration of response.

The pertuzumab+trastuzumab+docetaxel arm showed statistically significant improvement over placebo+trastuzumab+docetaxel arm with respect to IRF-assessed PFS in the intent-to-treat (ITT) population [hazard ratio=0.618, 95% confidence interval: (0.510, 0.749), log-rank test stratified by prior treatment status and region, two-sided p-value<0.0001]. The median PFS and its 95% confidence interval in pertuzumab+trastuzumab+docetaxel and placebo+trastuzumab+docetaxel arms were 18.5 months [95% CI: (14.6 months, 22.1 months)] and 12.4 months [95% CI: (10.4 months, 13.2 months)], respectively. Analysis of the primary endpoint PFS as assessed by IRF is shown in Table 3 and the Kaplan Meier plot is shown in Figure 2. The OS data were not mature at the time of PFS analysis and an interim OS analysis with 165 deaths (42.86% of total number of deaths required for the final analysis) did not show a statistically

significant difference between the two arms in the ITT population [hazard ratio=0.642, 95% confidence interval: (0.470, 0.877), stratified log-rank test, two-sided p-value=0.0053, O'Brien-Fleming boundary p=0.0012 based on the observed number of deaths]. The OS median in either arm was not reached. Analysis of OS is shown in Table 4 and the Kaplan Meier plot of OS is presented in Figure 3. The overall response rate was 80.2% in the experimental arm and 69.3% in the control arm based on the responses assessed by the IRF. The p-value for ORR is not interpretable because the first secondary endpoint OS in the hierarchical testing did not show statistical significance at the interim analysis and ORR was the second secondary endpoint.

2. INTRODUCTION

2.1. Overview

Breast cancer is the most common form of malignancy in women, with a global prevalence of more than 1.3 million patients and a mortality of approximately 450,000 deaths per year. Most breast cancers in the Western world (around 94%-95% of patients in the US and Europe) are diagnosed when the cancer is still confined to the breast, with or without loco-regional lymph node spread. At this stage, the disease is usually operable and can be treated with curative intent. However, around 20%-45% of patients experience relapse and those with metastatic or unresectable disease are generally incurable. Patients with metastatic disease have a median survival of around 24 months and a 5-year life expectancy of 18%-23% in the US and Europe.

The HER2 receptor (human epidermal growth factor receptor 2) has emerged as one of the most important targets for the treatment of breast cancer. HER2 is involved in regulating cell growth, survival and differentiation. Amplification and/or overexpression of HER2 occurs in around 15% to 20% of breast cancers and is a hallmark of the HER2-positive and luminal-B intrinsic sub-types of breast cancer. HER2 overexpression/amplification is associated with increased tumor aggressiveness, higher rates of recurrence, and increased mortality.

2.1.1. Background

Pertuzumab (rhuMAb 2C4) is a recombinant, humanized, immunoglobulin (Ig)G1κ monoclonal antibody, which targets the human epidermal growth factor receptor 2 (HER2, also known as c-erbB-2), a transmembrane glycoprotein with intrinsic tyrosine kinase activity. Pertuzumab is a HER2 dimerization inhibitor. By binding to the subdomain 2 epitope of the extracellular domain of HER2, it prevents heterodimerization of HER2 with other members of the HER family (HER1, HER3 and HER4). As a result, ligand-activated downstream signalling is blocked by pertuzumab. Pertuzumab is also capable of activating antibody-dependent cell-mediated cytotoxicity (ADCC).

2.1.2. Regulatory History

Pertuzumab is a new molecular entity (NME). This application is based on a single randomized Phase III, two supporting Phase II and several other Phase I and Phase II trials. The trials were conducted under IND 9,900. The End-of-Phase 2 meeting for the pivotal trial WO20698/TOC4129g was held on April 17, 2007. There was no Special Protocol Assessment for the pivotal trial. The Pre-BLA meeting was held on September 30, 2011.

2.1.3. Specific Studies Reviewed

This application is based on one Phase III trial WO20698/TOC4129g (CLEOPETRA), two supporting Phase II studies (WO20697 [NEOSPHERE] and BO17929), as well as a number of Phase I and II studies.

Study WO20697 was a randomized, multicenter, open-label Phase II study in patients with locally advanced, inflammatory or early stage HER2 positive breast cancer in neoadjuvant setting. The study had four arms: trastuzumab + docetaxel, trastuzumab + docetaxel + pertuzumab, trastuzumab + pertuzumab and pertuzumab + docetaxel. The primary endpoint was pathological complete response rate. Study BO17929 was a single-arm study of pertuzumab in combination with trastuzumab in patients with HER2-positive metastatic breast cancer.

This review is primarily based on the Phase III study. The CLEOPETRA study was a multicenter, international, randomized, double-blind, placebo-controlled, Phase III study to evaluate the efficacy of pertuzumab in combination with trastuzumab and docetaxel compared to placebo in combination with trastuzumab and docetaxel in patients with HER2-positive locally recurrent, unresectable or metastatic breast cancer who have not received anti-cancer treatment for their metastatic disease (except a maximum of one prior hormonal treatment for MBC).

Patients were randomized in a 1:1 ratio to receive either pertuzumab intravenously at a loading dose of 840 mg/kg followed by a dose of 420 mg/kg every 3 weeks or pertuzumab placebo intravenous infusion every 3 weeks. All patients received intravenous trastuzumab at a loading dose of 8 mg/kg followed by 6 mg/kg every 3 weeks and intravenous docetaxel at a dose of 75 mg/m^2 every 3 weeks for at least 6 cycles.

The randomization was stratified by prior treatment status (de novo vs. adjuvant or neoadjuvant) and region (Europe vs. North America vs. South America vs. Asia). The study started on February 12, 2008. The data cut-off date was May 13, 2011.

A total of 808 patients were randomized, 402 to the experimental arm and 406 to the control arm. Of the randomized patients, 2 were men and 806 were women, 480 were White, 261 were Asian and the median age was 54 years (age range: 22 to 89 years). Randomized patients were enrolled at 204 centers in 25 countries. There were 116 patients from US.

2.2. Data Sources

Data used for this review are from the electronic submission dated December 6, 2011. The path is \\Cbsap58\M\eCTD_Submissions\STN125409\0000\m5\datasets\wo20698-toc4129g.

3. STATISTICAL EVALUATION

3.1. Data and Analysis Quality

Overall the data and analysis quality of the submission was acceptable for the reviewer to be able to perform the statistical review.

3.2. Evaluation of Efficacy

The applicant has submitted efficacy results from one Phase III study (WO20698/TOC4129g [CLEOPETRA]) titled "A Phase III, randomized, double-blind, placebo-controlled clinical trial to evaluate the efficacy and safety of pertuzumab + trastuzumab + docetaxel vs. placebo + trastuzumab + docetaxel in previously untreated HER2-positive metastatic breast cancer." and two Phase II studies (WO20697 [NEOSPHERE] and BO17929). Study WO20697 was a randomized, multicenter, open-label Phase II study in patients with locally advanced, inflammatory or early stage HER2 positive breast cancer in neoadjuvant setting. The study had four arms: trastuzumab + docetaxel, trastuzumab + docetaxel + pertuzumab, trastuzumab + pertuzumab and pertuzumab + docetaxel. The primary endpoint was pathological complete response rate. Study BO17929 was a single-arm study of pertuzumab in combination with trastuzumab in patients with HER2-positive metastatic breast cancer. This review will be primarily based on the Phase III study.

3.2.1. Study Objectives

3.2.1.1. Primary Objective

The primary objective of Study WO20698/TOC4129g was to compare progression-free survival (PFS), based on tumor assessments by an independent review facility (IRF), between patients receiving pertuzumab in combination with trastuzumab and docetaxel and the patients receiving placebo with trastuzumab and docetaxel.

3.2.1.2. Secondary Objectives

The secondary objectives were to:

- Compare overall survival (OS) between the two treatment arms,
- Compare PFS between the two treatment arms based on investigator assessment of progression,
- Compare the overall objective response rate between the two treatment arms,
- Compare the duration of objective response between the two treatment arms,
- Compare the safety profile between the two treatment arms,
- Compare the time to symptom progression between the two treatment arms, as assessed by the Functional Assessment of Cancer Therapy (FACT) Trial Outcome Index—Physical/Functional/Breast (TOI-PFB),

- Evaluate if biomarkers from tumor tissues or blood samples (e.g., HER3 expression, Fcy-Receptor polymorphisms, and serum ECD/HER2 and/or HER ligand concentrations) correlate with clinical outcomes.

3.2.2. Study Design

This study was a multicenter, international, randomized, double-blind, placebo-controlled, Phase III study.

Patients were randomized in a 1:1 ratio to one of two treatment arms:

Arm A: Placebo + Trastuzumab +Docetaxel (Pla+T+D):

- Pertuzumab placebo: IV infusion every 3 weeks (q3w)
- Trastuzumab: Loading dose of 8 mg/kg IV, followed by 6 mg/kg IV q3w
- Docetaxel dose of 75 mg/m^2 IV q3w for at least six cycles

Arm B: Pertuzumab + Trastuzumab +Docetaxel (Ptz+T+D)

- Pertuzumab: Loading dose of 840 mg/kg IV, followed by 420 mg/kg IV q3w
- Trastuzumab: loading dose of 8 mg/kg IV, followed by 6 mg/kg IV q3w
- Docetaxel dose of 75 mg/m^2 IV q3w for at least six cycles

An Interactive Voice Response System (IVRS) was used for randomization which was stratified by prior treatment status (de novo vs. prior adjuvant or neoadjuvant therapy) and region (Europe vs. North America vs. South America vs. Asia). A complete block randomization scheme was applied to achieve balance in treatment assignment within each of the eight strata.

At the investigator's discretion, the docetaxel dose could be increased to 100 mg/m^2 for patients who tolerated at least one cycle without significant toxicities.

Treatment with pertuzumab/placebo and trastuzumab was to continue until investigator-assessed PD or unmanageable toxicity. Treatment with docetaxel was to continue for a minimum of six cycles, unless the patient experienced unacceptable toxicity or PD. After six cycles, continuation of docetaxel was at the discretion of the Investigator. If pertuzumab/placebo and/or trastuzumab had to be permanently discontinued or withheld for more than two cycles, the patient was taken off the study treatment. However, if docetaxel had to be permanently discontinued for reasons related to toxicity, the patient could continue with pertuzumab/placebo and trastuzumab. Patients did not receive open-label pertuzumab after discontinuation of study treatment and remained blinded to treatment allocation.

The study population for this trial comprised patients at least 18 years old with previously-untreated (in the metastatic setting), HER2-positive, metastatic or locally recurrent, unresectable breast cancer. This population included patients who had not been treated previously with chemotherapy and/or biologic therapy (including approved or investigational tyrosine kinase/ HER inhibitors or vaccines) for their metastatic disease. Patients were allowed prior adjuvant hormonal therapy and one line of hormonal therapy for metastatic disease. Patients with stage IV disease at initial disease presentation or PD occurring ≥ 12 months after neoadjuvant or adjuvant therapy were included. Trastuzumab and/or taxanes were acceptable neoadjuvant or adjuvant treatments.

Major inclusion criteria included histologically or cytologically confirmed adenocarcinoma of the breast with locally recurrent or metastatic disease, and candidate for chemotherapy, HER2-positive (defined as 3+ IHC or FISH amplification ratio ≥ 2.0) MBC confirmed by a Sponsor-designated central laboratory, LVEF ≥ 50% at baseline (within 42 days of randomization) as determined by either ECHO or MUGA and ECOG performance status 0 or 1.

3.2.3. Schedule of Assessments

Baseline tumor assessment was to be performed within 28 days prior to randomization. Tumor assessments (and assessments performed at the time of tumor assessments) to be performed until IRF-confirmation of PD. Tumor assessments were scheduled every 9 weeks ± 3 days from the date of randomization. If a tumor assessment was performed early or late, subsequent assessments were conducted according to the original schedule of every 9 weeks from the date of randomization. All patients had a minimum of a chest and abdomen CT scan. PET scans were not considered for assessments of efficacy at any time during the study (except as specified for bone scans in the absence of radioactive isotopes). Bone scans were preformed as clinically indicated. Tumor assessments were conducted until IRF-determined PD, even if treatment had been discontinued due of investigator-determined PD or unacceptable toxicity.

3.2.4. Efficacy Endpoints

Primary endpoint:

- Progression-free survival (PFS) based on tumor assessment by the IRF

Secondary endpoints include:

- Overall survival (OS)
- Progression-free survival (PFS) based on investigator assessment.
- Overall response rate (ORR)
- Duration of response (DR)

The primary endpoint PFS was defined as the time from randomization to the first documented PD, as determined by the IRF using RECIST version 1.0 or death from any

cause (within 18 weeks of last tumor assessment), whichever occurred first. Assessment of PD was based on a review of radiographic (MRI, CT, bone scans, chest x-ray, etc.), as well as cytologic (e.g. relevant cytology reports documenting malignant pleural effusions, bone marrow aspirations, cerebral spinal fluid, etc.), and photographic data, if available. The date of progression is assigned by the IRF as the earliest date of a scan where an overall disease status of PD is determined. If, for a given visit where a PD event is determined a patient had more than one method of imaging on varying dates, then the IRF was to assign the endpoint date as first date associated with that visit.

Patients who did not have documented IRF-determined PD at the time of the data cut-off for the primary efficacy analysis, or patients who have not died within 18 weeks of the last tumor assessment at which the IRF determined they were progression-free, were censored at the date of the last IRF-reviewed, evaluable tumor assessment. Patients who have died within 18 weeks of the last tumor assessments and who have no documented IRF-determined PD were included as having an event, with the event date for PFS as the date of death. This includes patients who have died within 18 weeks of the baseline tumor assessment, who have no post-baseline tumor assessment. If no tumor assessments were performed after the baseline visit, and the patient has not died within 18 weeks of the baseline tumor assessment, the patient was censored at Day 1.

Overall survival (OS) was defined as the time from the date of randomization to the date of death from any cause.

Patients who were alive or lost to follow up at the time of the analysis or who withdraw consent for survival follow-up were censored at the last known date. Patients with no post-baseline information were censored at Day 1.

Overall response was defined as a complete response [CR], or partial response [PR] determined by the IRF using RECIST on two consecutive occasions $\geq$ 4 weeks apart (patients without measurable disease or with disease localized only to the bone were not included in the analysis of objective response).

Duration of objective response, defined as the period from the date of initial confirmed partial or complete response until the date of PD or death from any cause (tumor responses were based on IRF evaluations using RECIST).

3.2.5. Sample Size Considerations

The sample size for this study was calculated based on PFS and OS. In order to detect a 33% improvement in median PFS (10.5 months in the control arm vs. 14 months in the experimental arm) or a hazard ratio of 0.75 at a two-sided significance level of 0.05 with 80% power and 1:1 randomization, a total of 381 IRF-assessed PFS events (corresponding to approximately 448 investigator assessed PFS events) were required. Based on the protocol-assumed accrual rate of approximately 40 patients per month after

a 9-month ramp-up period, it was estimated that 800 patients would need to be enrolled. The estimates for the 9-month ramp-up period detailed in the protocol are as follows:
Month 1 – one patient
Month 2 – two patients
Month 3 – four patients
Month 4 – eight patients
Month 5 – thirteen patients
Month 6 – nineteen patients
Month 7 – twenty-six patients
Month 8 – thirty-four patients
Month 9 – forty patients

The study was also adequately powered for detecting a 33% improvement in median OS (36 months in the control arm vs. 48 months in the experimental arm) or a hazard ratio of 0.75 with a 0.05 two sided level of significance and 80% power. This would require 385 deaths with one interim analysis at 50% events. The sample size was calculated such that approximately 50% deaths would be observed at the time of the final PFS analysis. It was estimated that the patients would need to be followed for an additional 29.5 months after completion of enrollment to observe 385 deaths.

3.2.6. Interim Analyses

No efficacy interim analysis of PFS was planned or conducted. A safety interim analysis of PFS was performed after the first 100 patients were followed for four months. The data monitoring committee (DMC) was unblinded for the interim safety analysis but all sponsor and safety management team personnel remained blinded. An independent cardiac review committee (CRC) was to review all suspected left ventricular systolic dysfunction (LVSD) in a blinded manner. The DMC could recommend stopping the study if, among the patients randomized at least 4 months prior to the data cut, the incidence of cardiac events (symptomatic LVSD events (deaths or non-deaths), non-LVSD cardiac deaths and probable cardiac deaths) based upon the CRC assessment was at least 9.3% higher in the pertuzumab arm compared with the control arm. The DMC could also recommend stopping the study if, in their opinion, the incidence of other clinically significant toxicities, such as neutropenia, neutropenic sepsis, or severe pulmonary toxicity, was unacceptably high in the pertuzumab arm compared with the control arm. The recommendation from the DMC following the safety interim analysis was to continue the study unchanged.

An interim efficacy analysis of OS was planned and conducted at the time of the final PFS analysis. It was estimated that there would be 50% OS events at the time of the interim analysis. A Lan-DeMets alpha-spending function with an O'Brien-Fleming stopping boundary was used for the interim analysis of OS.

According to the Lan-DeMets alpha spending function for O'Brien-Fleming boundary and the actual number of events, the alpha for the interim OS analysis with cut-off date

May 13, 2011 was 0.0012. At the interim OS analysis, there were 165 deaths (42.8% of the OS events needed for the final analysis). The observed p-value (0.0050) did not cross the stopping boundary ($p \leq 0.0012$).

3.2.7. Efficacy Analysis Methods

3.2.7.1. Analysis Populations

The Intent-to-Treat (ITT) population is defined as all randomized patients. All efficacy analyses are based on the intent-to treat (ITT) population with patients included under the treatment arm to which they were randomized. For certain outcomes, the analysis are based on sub-sets of the ITT population, as the outcome may only be relevant for specific patients. For objective response rate and time to response, only patients with measurable disease at baseline are included in the analysis. For duration of response, only responders are included in the analysis.

All patients randomized and have received any amount of study medication are included in the Safety Analysis Population. All safety analyses are based on this population with treatment assignment designated according to actual treatment received.

3.2.7.2. Analysis of Primary Endpoint

The primary endpoint was IRF-assessed PFS. A log-rank test, stratified by prior treatment status (de novo vs. prior adjuvant or neoadjuvant therapy) and region (Europe vs. North America vs. South America vs. Asia), was used to compare PFS between the two treatment arms.

The Kaplan-Meier approach was used to estimate median PFS for each treatment arm and the Cox proportional hazard model, stratified by prior treatment status and region was used to estimate the hazard ratio between the two treatment arms and its 95% confidence interval (CI).

The analyses were performed in pre-defined demographic subgroups as appropriate if there was a reasonable sample size. Univariate and multivariate Cox regression analyses were also performed to investigate the association between the pre-defined stratification and baseline prognostic covariates with PFS.

3.2.7.3. Analysis of Secondary Endpoints

Analysis methods of OS were the same as those of the primary endpoint.

PFS, based on investigator assessment, was a secondary efficacy endpoint. Data for patients without documented PD or who did not die within 18 weeks of the last tumor assessment were censored at the time of the last investigator tumor assessment (or, if no tumor assessments are performed after the baseline visit, at Day 1). Analysis methods for this endpoint were same as those for the primary endpoint.

Only patients with IRF-determined measurable disease at baseline were included in the analysis of the objective response. Objective response was based on the best overall response recorded from the start of trial treatment until IRF-assessed PD, death or first administration of next-line anti-cancer therapy (NACT), whichever occurs earliest. Patients without a post-baseline IRF-assessed tumor assessment were considered to be non-responders.

An estimate of the objective response rate and its 95% CI were calculated for each treatment arm. The Mantel-Haenszel χ^2 test stratified by prior treatment status and region was used to compare the objective response rate between the two treatment arms. An unadjusted Fisher's exact test result was provided as a sensitivity analysis. As a sensitivity analysis, Investigator-assessed objective response was evaluated, based on patients with Investigator-determined measurable disease at baseline.

Duration of objective response was based on IRF assessments. No formal hypothesis testing was performed on this endpoint, as the subgroup of patients with objective response is not a randomized subset. Median duration of objective response for each treatment arm was estimated using the Kaplan-Meier approach. The hazard ratio between the two treatment arms was estimated using Cox regression. As a sensitivity analysis, duration of objective response was repeated based on Investigator assessments.

The secondary endpoints were to be tested in the following order.

1. OS
2. ORR

Reviewer's comment:

Except for OS and ORR no Type I error rate has been adjusted for analysis of other secondary endpoints. Therefore, p-values for those secondary endpoints are not interpretable.

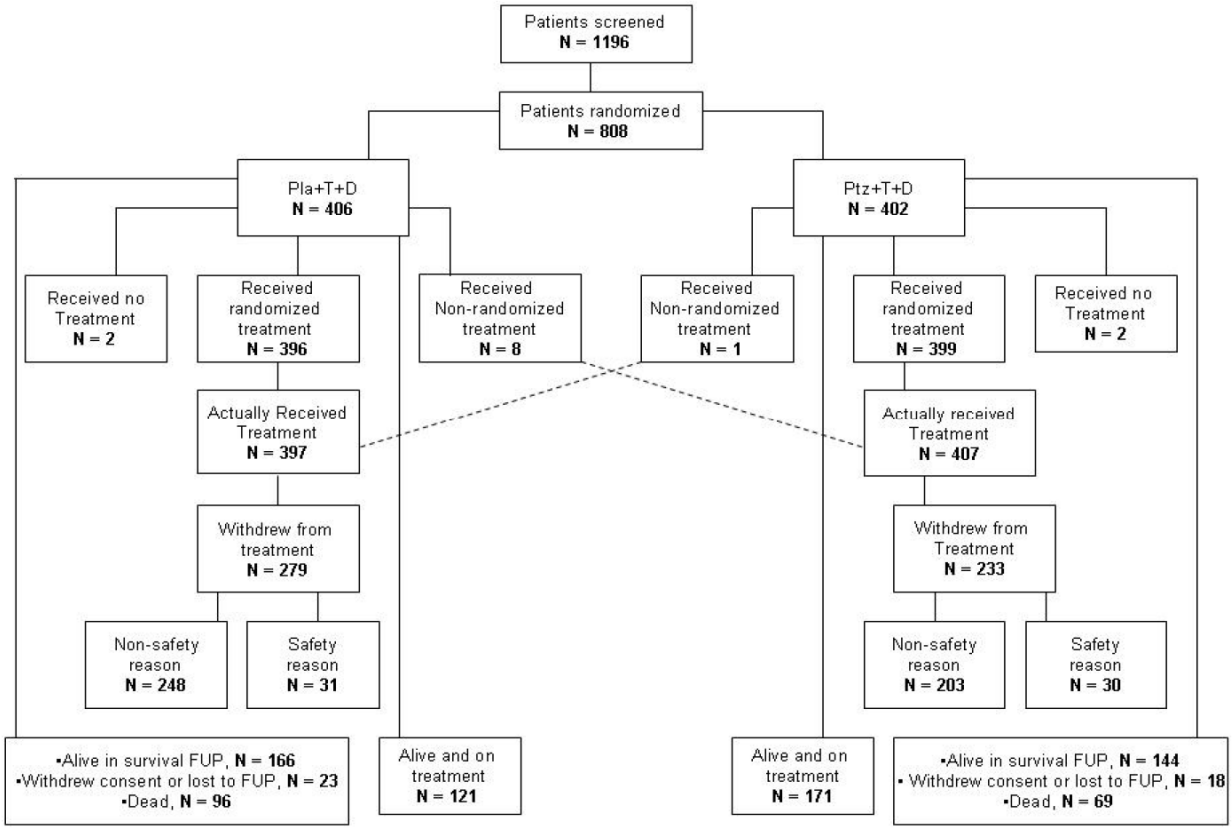
3.2.8. Sponsor's Results and FDA Statistical Reviewer's Findings/Comments

A total of 808 patients were randomized, 402 to the pertuzumab arm and 406 to the placebo arm. Two patients in each arm did not receive any study treatments. Patients were recruited at 204 centers in 25 countries. The data cut-off date for this study was May 13, 2011.

3.2.8.1. Patient Disposition

The patient disposition is presented in Figure 1.

Figure 1: Patient Disposition



Source: Clinical study report submitted in the BLA.

3.2.8.2. Baseline Characteristics

The treatment arms were well balanced with respect to general demographic characteristics (gender, race, age and geographic region) at baseline. Because of the disease almost all patients were female. Majority of the patients were White (59.4%) while the second largest race among the patients was Asian (32.3%). The mean and median age of patients at randomization was 53.49 and 54 years, respectively, with an overall age range of 22 to 89 years. Approximately 16% of patients were elderly. The study recruited patients from 25 countries. Approximately 37.9% patients were from Europe, 31.3% were from Asia and 14.4% patients were from US. A summary of demographic characteristics at baseline is presented in Table 1.

Baseline characteristics of the patients are presented in Table 2. Approximately 65% patients enrolled in this study had ECOG PS of 0, 35% had ECOG PS 1. The percentage of patients with ECOG performance status 0 was higher in the pertuzumab arm (68%) than in the placebo arm (61%). Approximately 47% patients previously received adjuvant or neo-adjuvant therapy, 78% patients had visceral disease at baseline, 48% were hormone receptor positive and 50% were negative.

Table 1: Demographic Characteristics: Gender, Race and Age and Geographic Region at Randomization in the ITT Population

		Pertuzumab + Trastuzumab + Docetaxel (N=402)	**Placebo + Trastuzumab + Docetaxel (N=406)**	**All (N=808)**
Gender	Female	402 (100.00%)	404 (99.51%)	806 (99.75%)
	Male	0 (0.00%)	2 (0.49%)	2 (0.25%)
Race	White	245 (60.95%)	235 (57.88%)	480 (59.41%)
	Black	10 (2.49%)	20 (4.93%)	30 (3.71%)
	Asian	128 (31.84%)	133 (32.76%)	261 (32.30%)
	Other	16 (3.98%)	14 (3.45%)	30 (3.71%)
Age Group in Years	<65	342 (85.07%)	339 (83.50%)	681 (84.28%)
	≥65	60 (14.93%)	67 (16.50%)	127 (15.72%)
Age in Years at Randomization	Mean, SD	53.44, 10.94	53.55, 11.35	53.49, 11.14
	Min, Max	22, 82	27, 89	22, 89
	Q1, Median, Q3	46, 54, 60	46, 54, 61	46, 54, 61
Geographic region	Asia	125 (31.09%)	128 (31.53%)	253 (31.31%)
	Europe	154 (38.31%)	152 (37.44%)	306 (37.87%)
	North America	67 (16.67%)	68 (16.75%)	135 (16.71%)
	South America	56 (13.93%)	58 (14.29%)	114 (14.11%)
US vs. Non-US	US	61 (15.17%)	55 (13.55%)	116 (14.36%)
	Non-US	341 (84.83%)	351 (86.45%)	692 (85.64%)

Table 2: Baseline Characteristics

		Pertuzumab + Trastuzumab + Docetaxel (N=402)		Placebo + Trastuzumab + Docetaxel (N=406)		All (N=808)	
Prior Treatment Status	Adjuvant or neo-adjuvant	184	(45.77%)	192	(47.29%)	376	(46.53%)
	De novo	218	(54.23%)	214	(52.71%)	432	(53.47%)
Visceral Disease Status	Visceral	314	(78.11%)	316	(77.83%)	630	(77.97%)
	Non-visceral	88	(21.89%)	90	(22.17%)	178	(22.03%)
Hormone Receptor (ER/PgR) Status	Positive	189	(47.01%)	199	(49.01%)	388	(48.02%)
	Negative	212	(52.74%)	196	(48.28%)	408	(50.50%)
	Unknown	1	(0.25%)	11	(2.71%)	12	(1.49%)
ECOG Performance Status	0	274	(68.16%)	248	(61.08%)	522	(64.60%)
	1	125	(31.09%)	157	(38.67%)	282	(34.90%)
	2	3	(0.75%)	0	(0.00%)	3	(0.37%)
	3	0	(0.00%)	1	(0.25%)	1	(0.12%)

3.2.8.3. Primary Efficacy Analysis

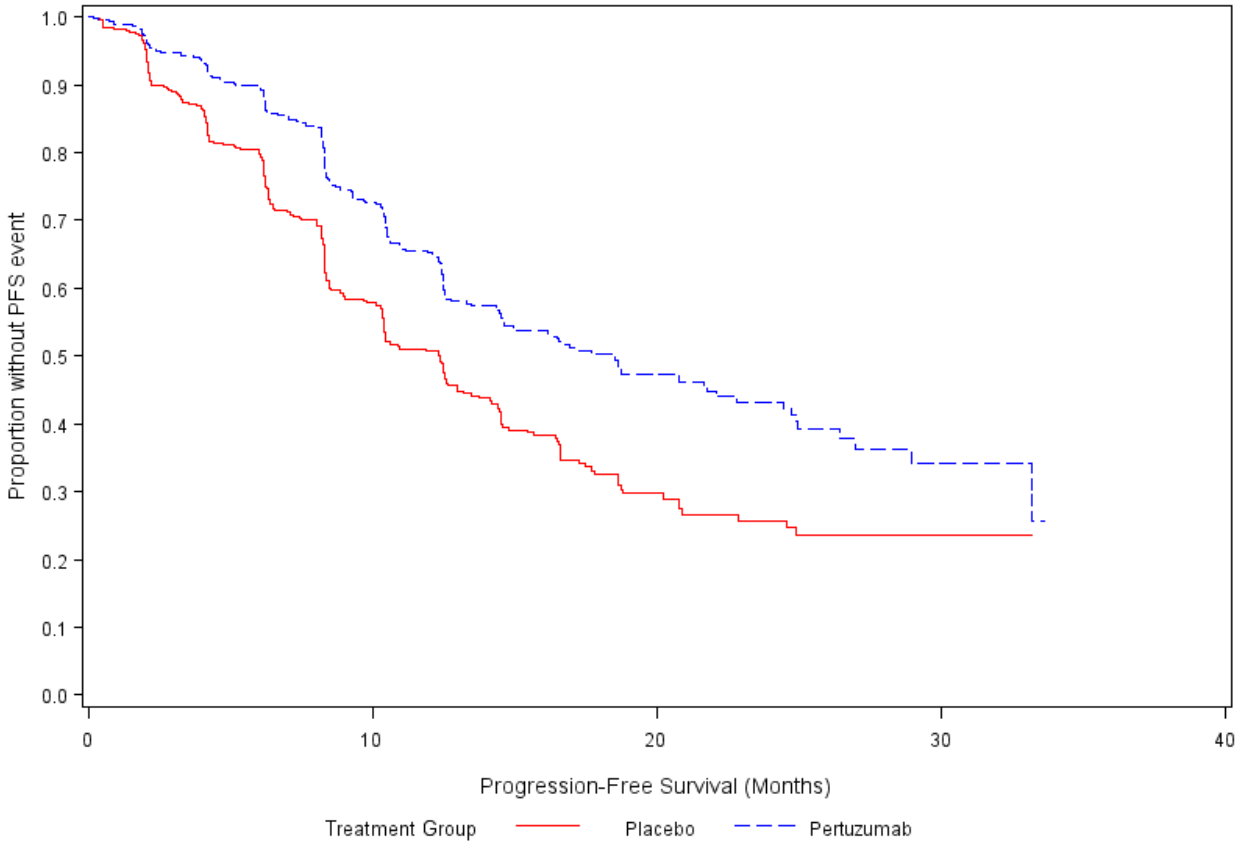
The primary efficacy analysis comparing progression-free survival (PFS) between Pertuzumab+trastuzumab+docetaxel and placebo+trastuzumab+docetaxel in the ITT population based on the independent radiology facility (IRF) assessment using log-rank test stratified by prior treatment status and region, the same stratification factors used at randomization, is presented in Table 3. The corresponding Kaplan-Meier plot is given in Figure 2. The PFS improvement in pertuzumab arm over placebo arm was statistically significant (stratified log-rank test, two-sided p-value < 0.0001). The PFS hazard ratio of pertuzumab over placebo using a stratified Cox model with the same stratification factors prior treatment status and region and its 95% confidence intervals were 0.618 [95% CI: (0.510, 0.749)]. The difference in median PFS between two arms is approximately 6 months.

Table 3: Analysis of PFS Based on IRF Assessment in the ITT Population

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)	P-value[3]
Pertuzumab*	402	191 (47.51%)	18.5 (14.6, 22.1)	0.618 (0.510, 0.749)	<0.0001
Placebo*	406	242 (59.61%)	12.4 (10.4, 13.2)		

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. [3]: Based on two-sided log-rank test, stratified by prior treatment status and region.

Figure 2: Kaplan-Meier Plot of PFS in the ITT Population Based on IRF Assessment



3.2.8.4. Secondary Efficacy Analyses

With this application, the applicant submitted the results of the interim OS analysis based on data cut-off of 13 May, 2011. The results of the interim OS analysis in the ITT population are presented in Table 4. The corresponding Kaplan-Meier plot is given in Figure 3. There was no statistically significant difference between trastuzumab+trastuzumab+docetaxel and placebo+trastuzumab+docetaxel arms with respect to OS at the interim analysis (log-rank test stratified by prior treatment status and region, nominal two-sided p-value 0.0053). The p-value did not cross the O'Brien-Fleming boundary which is 0.0012 with the observed number of 165 deaths (42.86% of total number of deaths required for the final OS analysis). The Kaplan-Meier plot of pertuzumab arm lies completely above that of placebo arm. The hazard ratio for OS based

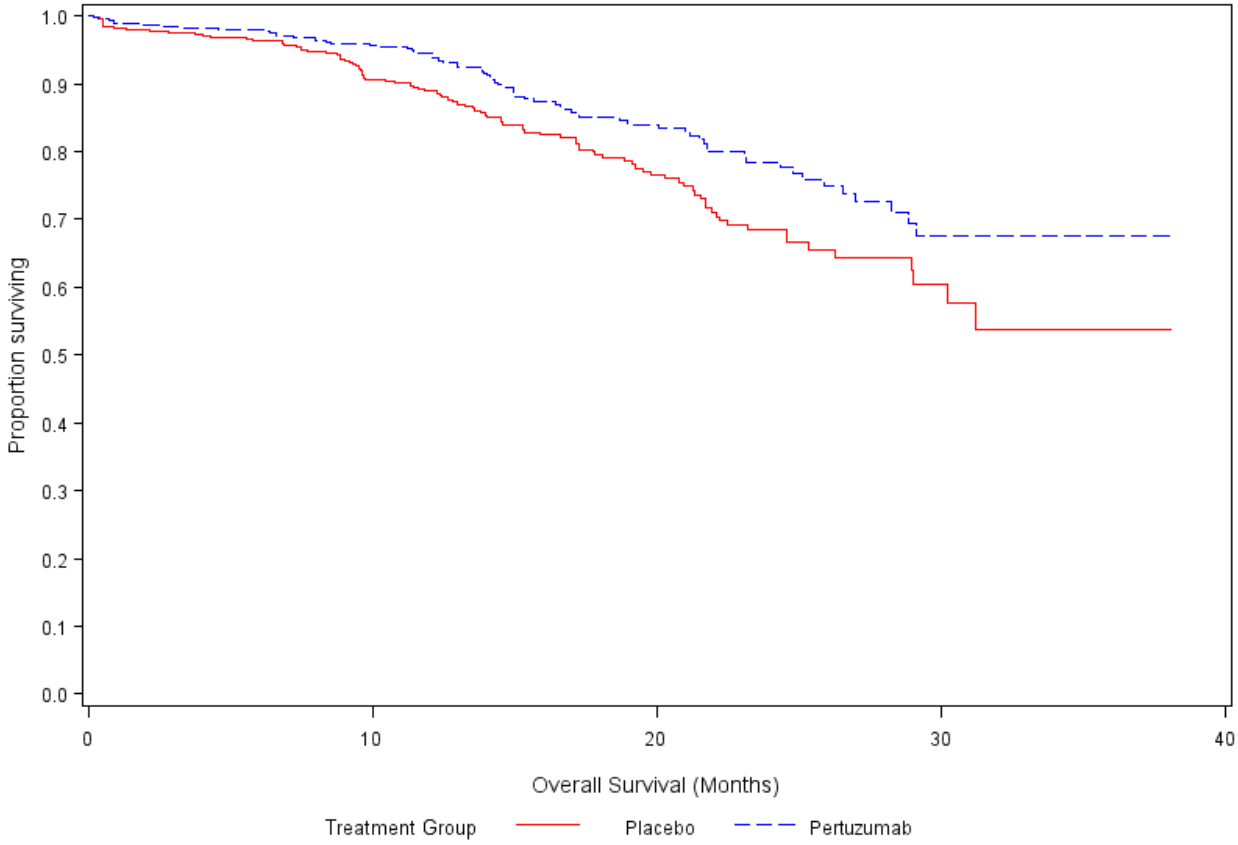
on the Cox model stratified by prior treatment status and region was 0.642 [95% CI: (0.470, 0.877)] at the interim analysis.

Table 4: Interim Analysis of OS in the ITT Population (May 13, 2011 Cut-off)

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)	P-value[3]
Pertuzumab*	402	69 (17.16%)	NE (NE, NE)	0.642 (0.470, 0.877)	0.0053
Placebo*	406	96 (23.65%)	NE (30.2, NE)		

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. [3]: Based on two-sided log-rank test, stratified by prior treatment status and region.

Figure 3: Kaplan-Meier Plot of OS in the ITT Population at the Interim Analysis



The ORR was the second secondary endpoint in the hierarchical testing. Table 5 shows the analysis of response rate based on IFR assessment. Only those patients with measurable disease at baseline were included in this analysis.

Table 5: Response Rate Based on IRF Assessment

	Randomization Group	
	Pertuzumab + Trastuzumab + Docetaxel (N=343)	**Placebo + Trastuzumab + Docetaxel (N=336)**
Complete Response	19 (5.54%)	14 (4.17%)
Partial Response	256 (74.64%)	219 (65.18%)
Objective Response Rate	80.2%	69.3%
P-value*	0.0011	

*: Based on two sided Mantel-Haenszel test stratified by prior treatment status and region

There were 19 complete responses and 256 partial responses in the pertuzumab + trastzumab + docetaxel arm and 14 complete responses and 219 partial responses in the placebo + trastzumab + docetaxel arm. The objective response rate was 80.2% in the pertuzumab + trastzumab + docetaxel arm and 69.3% in the placebo + trastzumab + docetaxel arm based on the responses assessed by the IRF. The P-value for Cochran-Mantel-Haenszel test stratified by prior treatment status and region is 0.0011.

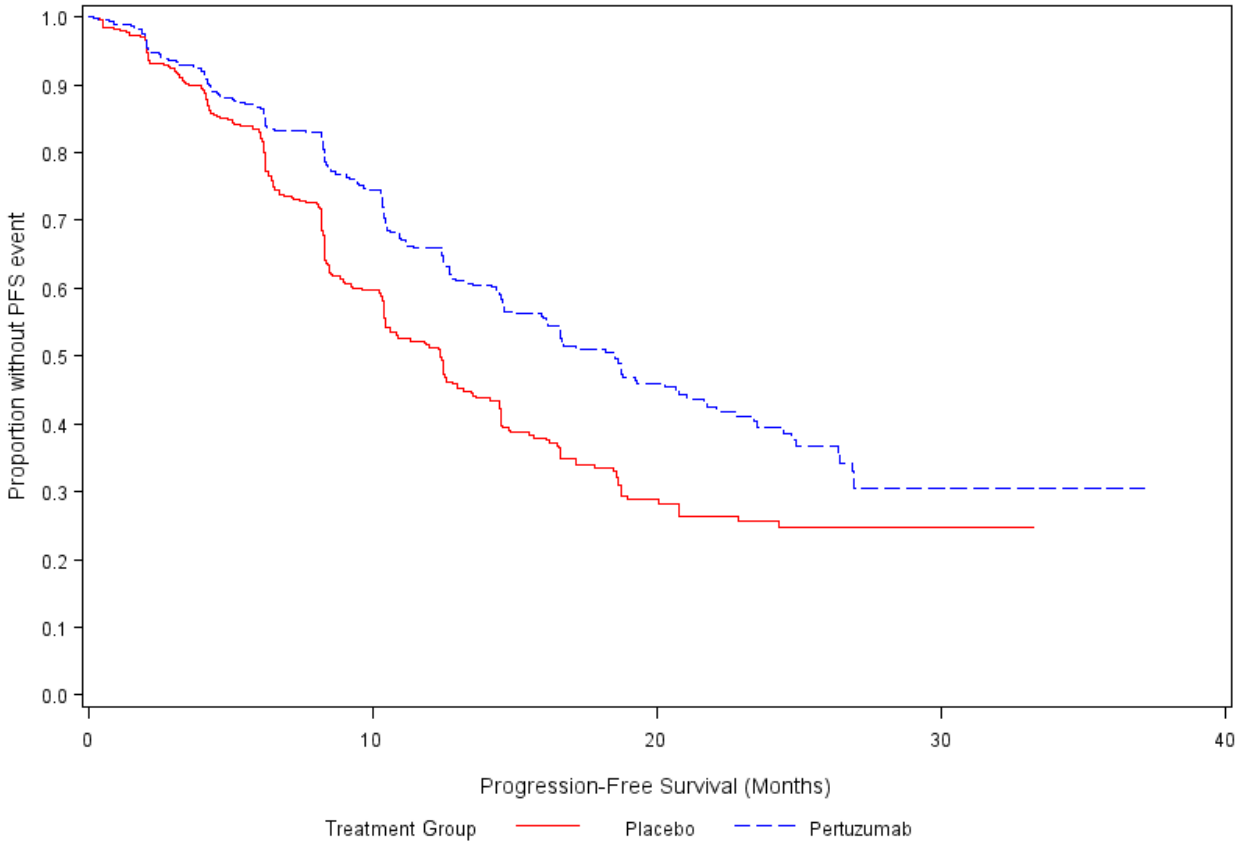
Investigator assessed PFS was a secondary endpoint. The analysis of PFS based on investigator assessment is presented in Table 6 and the Kaplan-Meier plot is presented in Figure 4. The results are similar to those for PFS based on IRF assessment.

Table 6: Analysis of PFS Based on Investigator's Assessment in the ITT population

Treatment	**Number of Patients**	**Number (%) Failed**	**Median in Months[1] (95% CI)**	**Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)**	**P-value[3]**
Pertuzumab*	402	201 (50.00%)	18.5 (16.1, 21.1)	0.651 (0.540, 0.785)	<0.0001
Placebo*	406	250 (61.58%)	12.4 (10.4, 13.2)		

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. [3]: Based on two-sided log-rank test, stratified by prior treatment status and region.

Figure 4: Kaplan-Meier Plot of PFS Based on Investigator's Assessment in the ITT Population



Reviewer's Comments:

1. Lan-DeMets alpha spending function has been used to adjust for multiple OS analyses. According to the Lan-DeMets alpha spending function for O'Brien-Fleming boundary and the actual number of events, the alpha for the interim OS analysis with cut-off date May 13, 2011 would be 0.0012 (calculated using East Version 5).
2. Except for OS and ORR, there was no adjustment in Type I error rate for other secondary endpoints. Therefore, p-values for secondary endpoints other than OS and ORR are not interpretable. The p-value for PFS based on investigator's assessments is not interpretable for that reason.
3. The p-value for the test of ORR is currently not interpretable because the previous endpoint OS in the hierarchical testing is not statistically significant at the interim analysis.

3.2.8.5. Sensitivity Analyses of PFS

Several sensitivity analyses of PFS are shown below.

1. Table 7 shows the PFS sensitivity analysis using the earlier of IRF-assessed and investigator assessed PD dates.
2. Table 8 shows the PFS sensitivity analysis by censoring PFS at the last IRF assessment before the next line anti-cancer therapy if the patient starts a new line of anti-cancer therapy before the IRF-assessed PD.
3. Table 9 shows the PFS sensitivity analysis for PFS events on treatment (no later than 42 days after the last treatment intake). If progression or death does not occur within this time window, PFS is censored at the last IRF assessment on treatment.
4. Table 10 shows the PFS sensitivity analysis where the expected date of the earliest missed assessment is used as the PFS event date if IRF-assessed PD or death occurs after one or more missed assessments.
5. Table 11 shows the PFS sensitivity analysis treating deaths occurring after 18 weeks of last IRF tumor assessment as a PFS event if there was no previous PFS event for that patient.
6. Table 12 shows the PFS sensitivity analysis where PFS is censored at the last IRF assessment before the discontinuation of treatment due to toxicity.
7. Table 13 shows the PFS sensitivity analysis with the censoring rules for missed assessments and new anticancer therapy as outlined in the FDA guidance for oncologic endpoints. If PD or death occurrs after two or more missed tumor IRF assessments, then PFS is censored at the last tumor assessment before the missed assessments. If a new anti-cancer therapy is started before a PFS event, PFS is censored at the last IRF assessment before the start of the new anti-cancer therapy.

Table 7: Sensitivity Analysis of PFS Using Earliest of PD dates Assessed by IRF and Investigator

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab[*] / Placebo[*] (95% CI)	P-value[3]
Pertuzumab[*]	402	226 (56.22%)	14.6 (12.5, 17.1)	0.663 (0.555, 0.792)	<0.0001
Placebo[*]	406	274 (67.49%)	10.4 (9.0, 12.4)		

[*]: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. [3]: Based on two-sided log-rank test stratified by prior treatment status and region.

Table 8: Sensitivity Analysis of PFS by Censoring PFS at the Time of the Next Line Anti-Cancer Therapy

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)	P-value[3]
Pertuzumab*	402	173 (43.03%)	18.7 (16.2, 24.7)	0.581 (0.475, 0.709)	<0.0001
Placebo*	406	231 (59.90%)	12.3 (10.4, 13.2)		

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. [3]: Based on two-sided log-rank test stratified by prior treatment status and region.

Table 9: Sensitivity Analysis of PFS on Treatment Based on IFR Assessments in the ITT Population

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)	P-value[3]
Pertuzumab*	402	165 (41.04%)	20.8 (16.5, 24.9)	0.581 (0.473, 0.714)	<0.0001
Placebo*	406	220 (54.19%)	12.4 (10.4, 14.2)		

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. [3]: Based on two-sided log-rank test stratified by prior treatment status and region.

Table 10: Sensitivity Analysis of PFS Treating PD after Missed Assessments to Have Occurred at the Earliest Missed Scheduled Assessment Date Based on IFR Assessments in the ITT Population

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)	P-value[3]
Pertuzumab*	402	191 (47.51%)	18.5 (14.5, 22.8)	0.621 (0.513, 0.753)	<0.0001
Placebo*	406	242 (59.61%)	12.3 (10.4, 12.5)		

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. [3]: Based on two-sided log-rank test stratified by prior treatment status and region.

Table 11: Sensitivity Analysis of PFS Including All Deaths as Events

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)	P-value[3]
Pertuzumab*	402	200 (49.75%)	17.2 (14.5, 21.6)	0.628 (0.520, 0.758)	<0.0001
Placebo*	406	251 (61.82%)	12.4 (10.4, 13.2)		

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. [3]: Based on two-sided log-rank test stratified by prior treatment status and region.

Table 12: Sensitivity Analysis of PFS Censoring PFS at the last IRF Assessment Before Treatment Discontinuation For Patients Discontinued Due to Toxicity

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab[*] / Placebo[*] (95% CI)	P-value[3]
Pertuzumab[*]	402	184 (45.77%)	18.5 (14.6, 22.8)	0.611 (0.502, 0.742)	<0.0001
Placebo[*]	406	237 (58.37%)	12.3 (10.4, 13.2)		

[*]: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. [3]: Based on two-sided log-rank test stratified by prior treatment status and region.

Table 13: Sensitivity Analysis of PFS with Censoring Rule According to FDA Guidance

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab[*] / Placebo[*] (95% CI)	P-value[3]
Pertuzumab[*]	402	173 (43.03%)	18.7 (16.2, 24.7)	0.581 (0.475, 0.709)	<0.0001
Placebo[*]	406	231 (56.90%)	12.3 (10.4, 13.2)		

[*]: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. [3]: Based on two-sided log-rank test, stratified by prior treatment status and region.

Reviewer's Comment:

The sensitivity analyses presented above show that the PFS results are robust under different assumptions. The hazard ratios are all very close to that of the primary PFS analysis and all P-values are less than 0.0001.

3.3. Evaluation of Safety

For a detailed safety evaluation, please refer to the clinical review of this application.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1. Gender, Race, Age and Geographic Region

Since all except two patients were females, no subgroup analysis was performed by gender. Efficacy by race was analyzed by exploratory analysis of PFS and is presented in Table 14. More than 91% patients were either White or Asian. Therefore, only those two racial groups are presented in Table 14. Efficacy by age group (<65 years, ≥65 years) was also analyzed by exploratory analysis of PFS and is presented in Table 15. Exploratory analysis of PFS by geographic region (Asia, Europe, North America and South America) is presented in Table 16 and the analyses of PFS for US and non-US subgroups are presented in Table 17. All PFS analyses in this section are based on IRF assessments. The reported hazard ratios are calculated using unstratified Cox models because subgroup sample sizes are smaller and a stratified analysis may lead to very small number of patients and events per cell.

Table 14: Exploratory Analysis of PFS by Race

Race	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab[*]/ Placebo[*] (95% CI)
White	Pertuzumab[*]	245	117 (47.76%)	17.2 (12.7, 22.8)	0.622 (0.487, 0.795)
	Placebo[*]	235	143 (60.85%)	10.4 (8.5, 12.5)	
Asian	Pertuzumab[*]	128	64 (50.00%)	18.6 (12.5, 24.9)	0.682 (0.489, 0.953)
	Placebo[*]	133	76 (57.14%)	13.2 (10.4, 16.4)	

[*]: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model.
NE: Not estimable.

Table 15: Exploratory Analysis of PFS by Age Group

Age Group	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab[*]/ Placebo[*] (95% CI)
<65 Years	Pertuzumab[*]	342	165 (48.25%)	17.2 (14.4, 22.1)	0.648 (0.528, 0.796)
	Placebo[*]	339	204 (60.18%)	12.5 (10.4, 13.8)	
≥ 65 Years	Pertuzumab[*]	60	26 (43.33%)	21.6 (12.4, NE)	0.517 (0.310, 0.861)
	Placebo[*]	67	38 (56.72%)	10.4 (8.3, 16.5)	

[*]: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model.
NE: Not estimable.

Table 16: Exploratory Analysis of PFS by Geographic Region

Geographic Region	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)
Asia	Pertuzumab*	125	62 (49.60%)	18.6 (12.5, 24.9)	0.677 (0.482, 0.951)
	Placebo*	128	73 (57.03%)	13.2 (10.4, 16.5)	
Europe	Pertuzumab*	154	80 (51.95%)	14.5 (11.9, 18.7)	0.721 (0.534, 0.973)
	Placebo*	152	92 (60.53%)	11.9 (8.5, 14.1)	
North America	Pertuzumab*	67	27 (40.30%)	22.8 (14.5, NE)	0.513 (0.315, 0.837)
	Placebo*	68	41 (60.29%)	10.4 (8.4, 14.5)	
South America	Pertuzumab*	56	22 (39.29%)	21.8 (12.4, NE)	0.456 (0.267, 0.778)
	Placebo*	58	36 (62.07%)	9.8 (6.5, 16.6)	

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. NE: Not estimable.

Table 17: Exploratory Analysis of PFS by US or NON-US

US or Non-US	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)
US	Pertuzumab*	61	25 (40.98%)	22.8 (14.4, NE)	0.494 (0.297, 0.822)
	Placebo*	55	37 (67.27%)	10.4 (8.3, 14.5)	
Non-US	Pertuzumab*	341	166 (48.68%)	16.9 (13.5, 21.6)	0.654 (0.532, 0.803)
	Placebo*	351	205 (58.40%)	12.4 (10.4, 14.1)	

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model. NE: Not estimable.

Reviewer's Comment:

Pertuzumab in combination with trastuzumab and docetaxel showed improvement over placebo in combination with trastuzumab and docetaxel across all age groups, race categories and geographic regions with respect to PFS but its improvement in the age group ≥ 65 years appears to be much more that than in the age group < 65 years and its improvement in North America and South America appears to be more than that in other geographical regions.

4.2. Other Special/Subgroup Populations

Exploratory analyses of PFS by ECOG performance status, prior treatment status (adjuvant or neoadjuvant therapy or de novo), visceral disease status (visceral or non-visceral disease),

hormone receptor (ER/PgR) status (positive or negative) and combination of ECOG PS and visceral disease status are presented in Table 18, Table 19, Table 20, Table 21 and Table 22, respectively. For 12 (1.49%) patients, the hormone receptor status was unknown. All PFS analyses are based on IRF assessments and hazards ratios were calculated using unstratified Cox models.

Table 18: Exploratory Analysis of PFS by Baseline ECOG Performance Status

ECOG Performance Status	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)
0	Pertuzumab*	274	123 (44.89%)	20.8 (14.6, 24.9)	0.701 (0.548, 0.897)
	Placebo*	248	131 (52.82%)	13.8 (10.9, 17.3)	
1	Pertuzumab*	125	67 (53.60%)	14.6 (12.5, 21.6)	0.555 (0.408, 0.754)
	Placebo*	157	110 (70.06%)	9.0 (8.3, 12.3)	

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model.

Table 19: Exploratory Analysis of PFS by Prior Treatment Status

Prior Treatment Status	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)
Adjuvant or Neo-adjuvant	Pertuzumab*	184	86 (46.74%)	18.6 (13.3, 27.0)	0.614 (0.465, 0.812)
	Placebo*	192	117 (60.94%)	12.4 (10.3, 14.5)	
De novo	Pertuzumab*	218	105 (48.17%)	17.2 (13.5, 22.8)	0.634 (0.488, 0.822)
	Placebo*	214	125 (58.41%)	12.4 (8.8, 14.2)	

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model.

Table 20: Exploratory Analysis of PFS by Visceral Disease Status

Visceral Disease Status	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab*/ Placebo* (95% CI)
Non-Visceral	Pertuzumab*	88	36 (40.91%)	20.8 (16.5, NE)	0.960 (0.607, 1.520)
	Placebo*	90	37 (41.11%)	17.3 (12.6, NE)	
Visceral	Pertuzumab*	314	155 (49.36%)	17.2 (13.3, 22.1)	0.554 (0.449, 0.684)
	Placebo*	316	205 (64.87%)	10.4 (8.5, 12.4)	

*: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model.
NE: Not estimable.

Table 21: Exploratory Analysis of PFS by Hormone Receptor (ER/PgR) Status

Hormone Receptor (ER/PgR) Status	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab[*] / Placebo[*] (95% CI)
Positive	Pertuzumab[*]	189	88 (46.56%)	17.2 (14.4, 27.0)	0.721 (0.546, 0.953)
	Placebo[*]	199	113 (56.78%)	14.4 (12.5, 16.6)	
Negative	Pertuzumab[*]	212	102 (48.11%)	18.7 (12.5, 24.9)	0.553 (0.424, 0.720)
	Placebo[*]	196	122 (62.24%)	8.9 (8.3, 10.4)	

[*]: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model.

Table 22: Exploratory Analysis of PFS by Baseline ECOG Performance Status and Visceral Disease Status

ECOG PS and Visceral Disease Status	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Pertuzumab[*] / Placebo[*] (95% CI)
ECOG PS 0 and Non-visceral Disease	Pertuzumab[*]	61	24 (39.34%)	20.8 (13.5, NE)	1.250 (0.690, 2.264)
	Placebo[*]	65	20 (30.77%)	NE (17.3, NE)	
ECOG PS 1 and Non-visceral Disease	Pertuzumab[*]	27	12 (44.44%)	16.7 (6.3, NE)	0.483 (0.222, 1.052)
	Placebo[*]	25	17 (68.00%)	12.5 (8.6, 15.5)	
ECOG PS 0 and Visceral Disease	Pertuzumab[*]	213	99 (46.48%)	18.5 (14.4, NE)	0.575 (0.438, 0.756)
	Placebo[*]	183	111 (60.66%)	12.3 (10.3, 14.1)	
ECOG PS 1 and Visceral Disease	Pertuzumab[*]	98	55 (56.12%)	14.3 (12.3, 21.6)	0.566 (0.405, 0.793)
	Placebo[*]	132	93 (70.45%)	8.3 (8.2, 10.5)	

[*]: Each treatment arm also contains trastuzumab and docetaxel. [1]: Kaplan-Meier estimate. [2]: Based on Cox model.
NE: Not estimable.

Reviewer's Comments:

1. All the subgroup analyses presented in this section are considered exploratory or hypothesis generating and no formal inference may be drawn.
2. The PFS improvement in the pertuzumab arm is consistent across subgroups defined by prior treatment status.
3. PFS improvement in pertuzumab appears higher in the subgroup of patients with baseline ECOG performance status 1 than in the subgroup with baseline ECOG performance status 0. It is also higher in the hormone receptor (ER/PgR) negative patients than in the hormone receptor positive patients.

4. Pertuzumab appears to have almost no improvement in PFS in the subgroup of patients with non-visceral disease. The sample size in that subgroup is not large, only 178.
5. The point estimate of the PFS hazard ratio of pertuzumab to placebo was 1.25 in the subgroup of patients with baseline ECOG performance status 0 and non-visceral disease indicating a possible trend against pertuzumab in that subgroup. However, the confidence interval for the PFS hazard ratio included the value 1. PFS improvements in the pertuzumab arm in other three subgroups defined by the combination of baseline ECOG performance status and visceral disease status are similar to or better than that in the overall population.

5. SUMMARY AND CONCLUSIONS

This application is based on one Phase III trial WO20698/TOC4129g (CLEOPETRA), two supporting Phase II studies (WO20697 [NEOSPHERE] and BO17929), as well as a number of Phase I and II studies. This review is primarily based on the Phase III study. The CLEOPETRA study was a multicenter, international, randomized, double-blind, placebo-controlled, Phase III study to evaluate the efficacy of pertuzumab in combination with trastuzumab and docetaxel compared to placebo in combination with trastuzumab and docetaxel in patients with HER2-positive locally recurrent, unresectable or metastatic breast cancer who have not received anti-cancer treatment for their metastatic disease (except a maximum of one prior hormonal treatment for MBC). Patients were randomized in a 1:1 ratio to receive either pertuzumab intravenously at a loading dose of 840 mg/kg followed by a dose of 420 mg/kg every 3 weeks or pertuzumab placebo intravenous infusion every 3 weeks. All patients received intravenous trastuzumab at a loading dose of 8 mg/kg followed by 6 mg/kg every 3 weeks and intravenous docetaxel at a dose of 75 mg/m^2 every 3 weeks for at least 6 cycles. The randomization was stratified by prior treatment status (de novo vs. adjuvant or neoadjuvant) and region (Europe vs. North America vs. South America vs. Asia). The study started on February 12, 2008. The data cut-off date was May 13, 2011. A total of 808 patients were randomized, 402 to the experimental arm and 406 to the control arm. Patients were enrolled at 204 centers in 25 countries. There were 116 patients from US. The primary efficacy endpoint was progression-free survival (PFS) as assessed by an independent radiology facility (IRF). The secondary efficacy endpoints included overall survival (OS), investigator-assessed PFS, overall response rate (ORR) as assessed by the IRF and duration of response.

The pertuzumab+trastuzumab+docetaxel arm showed statistically significant improvement over placebo+trastuzumab+docetaxel arm with respect to IRF-assessed PFS in the intent-to-treat (ITT) population [hazard ratio=0.618, 95% confidence interval: (0.510, 0.749), log-rank test stratified by prior treatment status and region, two-sided p-value<0.0001]. The OS data were not mature at the time of PFS analysis and an interim OS analysis with 165 deaths (42.86% of total number of deaths required for the final analysis) did not show a statistically significant difference between the two arms in the ITT population [hazard ratio=0.642, 95% confidence interval: (0.470, 0.877), stratified log-rank test, two-sided p-value=0.0053, O'Brien-Fleming boundary p=0.0012 based on the observed number of deaths]. The OS median in either arm was not reached. The overall response rate was 80.2% in the experimental arm and 69.3% in the control arm based on the responses assessed by the IRF. The p-value for ORR is not interpretable because the first secondary endpoint OS in the hierarchical testing did not show statistical significance at the interim analysis and ORR was the second secondary endpoint.

5.1. Statistical Issues and Collective Evidence

1. Except for OS and ORR, there was no adjustment in Type I error rate for other secondary endpoints. Therefore, p-values for secondary endpoints other than OS and ORR are not interpretable. The p-value for PFS based on investigator's assessments is not interpretable for that reason.

2. The p-value for the test of ORR is currently not interpretable because the previous endpoint OS in the hierarchical testing is not statistically significant at the interim analysis.

3. The sensitivity analyses show that the PFS results are robust under different assumptions. The hazard ratios are all very close to that of the primary PFS analysis and all P-values are less than 0.0001.

4. Pertuzumab in combination with trastuzumab and docetaxel showed improvement over placebo in combination with trastuzumab and docetaxel across all age groups, race categories, geographic regions and prior treatment status with respect to PFS.

5. PFS improvement in pertuzumab appears higher in the subgroup of patients with baseline ECOG performance status 1 than in the subgroup with baseline ECOG performance status 0. It also appears to be higher in the hormone receptor (ER/PgR) negative patients than in the hormone receptor positive patients, in the age group ≥ 65 years than in the age group < 65 years and in North America and South America than in other regions.

6. Pertuzumab appears to have almost no improvement in PFS in the subgroup of patients with non-visceral disease. The sample size in that subgroup is 178.

7. The point estimate of the PFS hazard ratio of pertuzumab to placebo was 1.25 in the subgroup of patients with baseline ECOG performance status 0 and non-visceral disease indicating a possible trend against pertuzumab in that subgroup. However, the confidence interval for the PFS hazard ratio included the value 1. PFS improvements in the pertuzumab arm in other three subgroups defined by the combination of baseline ECOG performance status and visceral disease status are similar to or better than that in the overall population.

5.2. Conclusions and Recommendations

The applicant has submitted results from one multicenter, phase III, randomized, double-blind clinical trial (Study WO20698/TOC4129g or CLEOPETRA) comparing pertuzumab, a new molecular entity (NME) in combination with trastuzumab and docetaxel, to placebo in combination with trastuzumab and docetaxel in patients with HER2-positive locally recurrent, unresectable or metastatic breast cancer who have not received anti-cancer treatment for their metastatic disease (except a maximum of one prior hormonal treatment for MBC). The pertuzumab arm showed statistically significant improvement over the placebo arm in progression-free survival (PFS) as assessed by an independent radiology facility in all randomized patients. At the time of the analysis of PFS, the overall survival (OS) data were not mature and the pertuzumab arm did not show statistically significant improvement with respect to overall survival (OS) in an interim analysis. The statistical results provide adequate evidence to support the PFS claim proposed in the BLA.

SIGNATURES/DISTRIBUTION LIST

Primary Statistical Reviewer: Somesh Chattopadhyay, Ph.D.
Date: May 10, 2012

Concurring Reviewer(s): Shenghui Tang, Ph.D., Team Leader
Thomas Gwise, Ph.D., Deputy Director

Statistical Team Leader: Shenghui Tang, Ph.D.

Biometrics Division Director: Rajeshwari Sridhara, Ph.D.

cc:
HFD-150/Ms. Amy Tilley
HFD-150/Dr. Gideon Blumenthal
HFD-150/Dr. Nancy Scher
HFD-150/Dr. Patricia Cortazar
HFD-711/Dr. Somesh Chattopadhyay
HFD-711/Dr. Shenghui Tang
HFD-711/Dr. Thomas Gwise
HFD-711/Dr. Rajeshwari Sridhara
HFD-700/Ms. Lillian Patrician

CHECK LIST

Number of Pivotal Studies: 1

Trial Specification

Protocol Number (s): WO20698C/TOC4129G
Protocol Title (optional): A Phase III, randomized, double-blind, placebo-controlled clinical trial to evaluate the efficacy and safety of pertuzumab + trastuzumab + docetaxel vs. placebo + trastuzumab + docetaxel in previously untreated HER2-positive metastatic breast cancer.
Phase: 3
Control: Placebo Control
Blinding: Double-Blind
Number of Centers: 204
Region(s) (Country): 25 countries (Brazil, Canada, China, Costa Rica, Croatia, Ecuador, France, Finland, Germany, Great Britain, Guatemala, Italy, Japan, Latvia, Macedonia, Mexico, Poland, Republic of Argentina, Republic of Korea, Republic of the Philippines, Russia, Singapore, Spain, Thailand, USA).
Duration: 12 February 2008 – 13 May 2011
Treatment Arms: Experimental: pertuzumab + trastuzumab + docetaxel, Control: placebo+ trastuzumab + docetaxel
Treatment Schedule:

Arm A (Placebo + Trastuzumab +Docetaxel):

- Pertuzumab placebo: IV infusion every 3 weeks (q3w)
- Trastuzumab: Loading dose of 8 mg/kg IV, followed by 6 mg/kg IV q3w
- Docetaxel dose of 75 mg/m^2 IV q3w for at least six cycles

Arm B (Pertuzumab + Trastuzumab +Docetaxel):

- Pertuzumab: Loading dose of 840 mg/kg IV, followed by 420 mg/kg IV q3w
- Trastuzumab: loading dose of 8 mg/kg IV, followed by 6 mg/kg IV q3w
- Docetaxel dose of 75 mg/m^2 IV q3w for at least six cycles

Randomization: Yes
Ratio: 1:1
Method of Randomization: Complete block randomization with stratification using IVRS
Stratification Factors: Prior treatment status (de novo vs. adjuvant or neoadjuvant) and region (Europe, North America, South America and Asia)
Primary Endpoint: Progression-free survival (PFS) based on tumor assessment by an independent review facility
Primary Analysis Population: ITT
Statistical Design: Superiority
Adaptive Design: No
Primary Statistical Methodology: Stratified log-rank test
Interim Analysis: Yes
If yes:
No. of Times: 1 for OS efficacy at the time of final PFS analysis

Method: O'Brien-Fleming

α Adjustment: Yes

α Spending Function: Lan-DeMets alpha spending function for O'Brien-Fleming boundary

DSMB: Yes

Sample Size: 808 (planned 800)

Sample Size Determination: Was it calculated based on the primary endpoint variable and the analysis being used for the primary variable? Yes

Statistic = Log-rank

Power= 80%

Δ= Hazard ratio of (median PFS of 14 months in experimental arm vs. 10.5 months in control arm

α = two-sided 0.05

- Was there an **Alternative Analysis** in case of violation of assumption; e.g., Lack of normality, Proportional Hazards Assumption violation. NA
- Were there any major changes, such as changing the statistical analysis methodology or changing the primary endpoint variable? No
- Were the **Covariates** pre-specified in the protocol? Yes.
- Did the Applicant perform **Sensitivity Analyses**? Yes.
- How were the **Missing Data** handled? For PFS and OS censoring was used.
- Was there a **Multiplicity** involved? Yes.

  If yes,

  Multiple Arms (Yes/No)? No.

  Multiple Endpoints (Yes/No)? Yes.

  Which method was used to control for type I error? Hierarchical testing for multiple secondary endpoints.
- **Multiple Secondary Endpoints**: Are they being included in the label? If yes, method to control for type 1 error. Yes. Hierarchical testing. The first secondary endpoint OS is not significant; the next secondary endpoint included descriptively.

Were Subgroup Analyses Performed (Yes/No)? Yes.

- Were there any **Discrepancies** between the protocol/statistical analysis plan vs. the study report? No.
- Overall, was the study positive (Yes/No)? Yes.

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

/s/

SOMESH CHATTOPADHYAY
05/10/2012

SHENGHUI TANG
05/10/2012

THOMAS E GWISE
05/10/2012

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

BLA Number: 125409 **Applicant: Genentech, Inc.** **Stamp Date: December 8, 2011**

Drug Name: Pertuzumab **NDA/BLA Type: Original BLA**

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

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

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? ___Yes_____

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

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

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

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Somesh Chattopadhyay	January 19, 2012
Reviewing Statistician	Date
Shenghui Tang	January 19, 2012
Supervisor/Team Leader	Date

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

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

/s/

SOMESH CHATTOPADHYAY
05/10/2012

SHENGHUI TANG
05/10/2012