Protocol MYL-Her 3001

A MULTICENTER, DOUBLE-BLIND, RANDOMIZED, PARALLEL-GROUP, PHASE III STUDY OF THE EFFICACY AND SAFETY OF HERCULES (MYLAN TRASTUZUMAB) PLUS TAXANE VERSUS HERCEPTIN® PLUS TAXANE AS FIRST LINE THERAPY IN PATIENTS WITH HER2-POSITIVE METASTATIC BREAST CANCER

Statistical Analysis Plan (SAP)

Version: 2.2

Author: Jinyu Yuan, Ph.D.

Date: 19 August 2016
Signature Page for Statistical Analysis Plan (SAP)

Approval of SAP– Version 2.2

Author:

________________________________________________________________________  __________
Jinyu Yuan, Ph.D.
Director, Biostatistics

Approval:

________________________________________________________________________  __________
Ashwani Marwah
Associate Scientific Manager, Biostatistics
Biocon Research Limited

________________________________________________________________________  __________
Eduardo Pennella, MD, MBA
Senior Director, Oncology/Hematology
Mylan Inc

________________________________________________________________________  __________
Hans-Friedrich Koch, Ph.D.
Senior Director, Biometrics
Mylan Inc

This Document is Company Confidential.
CONTENTS

ABBREVIATIONS ............................................................................................................. 5

1 AMENDMENTS FROM PREVIOUS VERSION(S) .................................................. 6

2 INTRODUCTION ........................................................................................................... 6

2.1 STUDY DESIGN ....................................................................................................... 6
  2.1.1 Sample Size Determination ............................................................................. 7

2.2 REPORTING OF RESULTS ...................................................................................... 7

2.3 STUDY OBJECTIVES ............................................................................................... 8
  2.2.1 Primary Objective ............................................................................................. 8
  2.2.2. Secondary Objectives .................................................................................... 8
  2.2.3 Exploratory Objectives .................................................................................... 9

3 TIMING OF ANALYSES AND UNBLINDING ......................................................... 9

3.1 INTERIM ANALYSIS .............................................................................................. 9

3.2 TIMING OF PART 1, PART 2, AND FINAL ANALYSIS OF OS .......................... 10
  3.2.1 Part 1: Timing and Data Cut-off ..................................................................... 10
  3.2.2 Part 2: Timing and Data Cut-off ..................................................................... 10
  3.2.3 Timing of Final Assessment of Overall Survival (36M/240D) ....................... 11

3.3 UNBLINDING ......................................................................................................... 11

4 HYPOTHESES AND DECISION RULES FOR THE PRIMARY ANALYSIS .... 11

4.1 STATISTICAL HYPOTHESES .............................................................................. 11

4.2 STATISTICAL DECISION RULES ....................................................................... 11

5 ANALYSIS SETS ....................................................................................................... 12

5.1 INTENT-TO-TREAT POPULATION #1 (ITT1) ................................................... 12

5.2 INTENT-TO-TREAT POPULATION #2 (ITT2) ................................................... 12

5.3 PER-PROTOCOL (PP) POPULATION ................................................................. 12

5.4 SAFETY (SAF) POPULATION .............................................................................. 13

5.5 PHARMACOKINETIC (PK) POPULATION .......................................................... 13

5.6 TREATMENT MISALLOCATIONS ....................................................................... 13

5.7 PROTOCOL DEVIATIONS ...................................................................................... 13

6 ENDPOINTS AND COVARIATES ..................................................................... 13

6.1 EFFICACY ENDPOINTS ....................................................................................... 13
  6.1.2 Secondary Endpoints ...................................................................................... 14

6.2 SAFETY ENDPOINTS ........................................................................................... 15
  6.2.1 Adverse Events ............................................................................................... 15
  6.2.2 Laboratory Safety Assessments ....................................................................... 15
  6.2.3 Other Safety Assessments .............................................................................. 15

6.3 POPULATION PHARMACOKINETIC ENDPOINTS ......................................... 16

6.4 COVARIATES AND STRATIFICATION FACTORS ......................................... 16
  6.4.1 Stratification Factors ....................................................................................... 16

This Document is Company Confidential.
7 HANDLING OF MISSING VALUES

8 STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES

8.1 ANALYSIS SOFTWARE

8.2 INTERIM ANALYSIS

8.2.1 Blinded Sample Size Re-estimation

8.2.2 Futility Analysis

8.3 DESCRIPTIVE SUMMARIES

8.4 PRIMARY ANALYSIS

8.5 SECONDARY ANALYSES

8.6 SENSITIVITY ANALYSES

8.6.1 Sensitivity Analyses for ORR

8.6.2 Sensitivity Analyses for TTP

8.6.3 Sensitivity Analyses for PFS

8.6.4 Sensitivity Analyses for OS

8.7 EXPLORATORY ANALYSIS

8.8 ANALYSIS OF OTHER ENDPOINTS

8.9 SAFETY ANALYSES

8.9.1 Adverse Events

8.9.2 Clinical Laboratory Data

8.9.3 Vital Signs

8.9.4 Electrocardiograms

8.9.5 Left Ventricular Ejection Fraction

8.10 ECOG PS

8.11 PK

8.12 IMMUNOCOMPATIBILITY

8.13 MULTICULTURALITY ADJUSTMENTS

8.14 THE INTERPRETATION OF PART 2 ANALYSIS RESULTS

9 DUAL SCOPE OF SAP

10 REFERENCES

Appendix A

Estimation of the Treatment Effect by Fixed-Effects Meta-Analysis of Historical Herceptin Studies and Justification of the Equivalence Region for MYL-HER 3001
ABBREVIATIONS

36M/240D The study timepoint when 36 months have passed since the last patient was randomized into the study or the 240th death has occurred, whichever occurs first.
AE Adverse Event
BOR Best Overall Response
BSSR Blinded Sample Size Re-estimation
CI Confidence Interval
CR Complete Response
CSR Clinical Study Report
CTC AE Common Terminology Criteria for Adverse Events
DR Duration of Response
DSMB Data and Safety Monitoring Board
ECD Extracellular Domain (of HER2)
ER/PgR Estrogen Receptor/Progesterone Receptor
FISH Fluorescent in situ hybridization
HER2 Human Epidermal Growth Factor Receptor 2
ITT Intent to Treat
MBC Metastatic Breast Cancer
NCI National Cancer Institute
OR Objective Response, Overall Response
ORR Objective Response Rate, Overall Response Rate
OS Overall Survival
PFS Progression Free Survival
PP Per Protocol
PR Partial Response
RECIST Response Evaluation Criteria in Solid Tumors
SAE Serious Adverse Event
SAF Safety Analysis Population
SAP Statistical Analysis Plan
SD Stable Disease; Standard Deviation
TEAE Treatment Emergent Adverse Event
TOST Two One-Sided Tests
TTP Time to Tumor Progression
UA Unable to Assess
1 AMENDMENTS FROM PREVIOUS VERSION(S)

Version 2.2 of the SAP amends Version 2.1 (Effective Date: 26 February 2016). It is clarified that three CSRs will be generated for this study. More details are provided for Part 2 safety analyses. One sensitivity analysis is added for PFS.

2 INTRODUCTION

This document describes the planned statistical analyses for Protocol MYL-Her 3001. Deviations from this analysis plan will be described in the Clinical Study Report.

This SAP uses the term “Mylan Trastuzumab” instead of “Hercules” when referring to the test treatment.

2.1 STUDY DESIGN

This is a multicenter, double-blind, randomized, parallel-group, Phase III study to compare the efficacy and safety of Mylan Trastuzumab plus docetaxel or paclitaxel (i.e., taxane) versus Herceptin® plus taxane in patients with HER2-positive metastatic breast cancer (MBC) with continuation of single agent Mylan Trastuzumab versus Herceptin® for patients who have at least obtained Stable Disease in order to evaluate continued safety and immunogenicity.

Eligible patients will be centrally randomized in a 1:1 fashion to receive either Mylan Trastuzumab in combination with taxane or Herceptin® in combination with taxane. Patients will be stratified by:

- Tumor progression into metastatic phase ≥2 years OR <2 years after primary diagnosis (calculated as time from primary tumor surgery until randomization). Patients diagnosed with primary metastatic disease will be classified together with the patients who progressed < 2 years.
- Estrogen receptor (ER)/progesterone receptor (PgR) status (ER and/or PgR positive/ER and PgR negative).
- Type of taxane received (e.g., paclitaxel or docetaxel). The choice of taxane will be the Investigator’s decision at the site level and will be determined before screening of the first patient at the site.

The study will consist of two parts.

In Part 1, Mylan Trastuzumab plus taxane or Herceptin® plus taxane will be administered for a minimum of 8 treatment cycles (1 treatment cycle = 3 weeks) unless the patient experiences unacceptable side effects, disease progression or is prematurely withdrawn from treatment. Patients will be monitored for response to therapy (stable disease [SD], partial response [PR] or complete response [CR]).
The primary analysis of the best ORR will be performed at Week 24. A formal interim analysis for futility will be conducted when at least 123 patients complete the Week 24 assessment in Part 1 and will be overseen by the data and safety monitoring board (DSMB).

Patients entering Part 2 will continue with either Mylan Trastuzumab alone or Herceptin® alone (i.e., without taxane) until disease progression, unacceptable toxicity or death, whichever occurs first, and will be followed for overall survival evaluation.

2.1.1 Sample Size Determination

Per FDA requirement, the equivalence analysis should be based on the ratio of ORRs. The equivalence region was justified by performing a fixed-effects meta-analysis with historical randomized Herceptin® trials to estimate the treatment effect of Herceptin plus chemotherapy versus chemotherapy alone. A description of this process is provided in Appendix A. A sample size of 410 patients (205 per treatment group) is required to provide at least 80% power to declare Mylan Trastuzumab equivalent to Herceptin® in the analysis of ORR at Week 24. This sample size assumes that both treatment groups will exhibit an ORR of 69% at Week 24 and that the ratio of Mylan Trastuzumab to Herceptin® will be analyzed with a two-sided 90% CI. If the 90% CI falls wholly within an equivalence region defined as (0.81, 1.24), then equivalence will be declared.

EMA has requested that the equivalence region is based on the difference of ORRs and defined as (-15%, 15%). A sample size of 400 patients (200 per treatment group) is required to provide 80% power to declare Mylan Trastuzumab equivalent to Herceptin® in the analysis of ORR at Week 24. This sample size assumes that both treatment groups will exhibit an ORR of 69% at Week 24. The difference of Mylan Trastuzumab and Herceptin® will be analyzed with a two-sided 95% CI. If the 95% CI falls wholly within the equivalence region (-15%, 15%), then equivalence can be declared.

Therefore, the sample size of 410 was chosen for the study to satisfy the requirement of both regulatory agencies for equivalence analysis. To arrive at the planned number of patients, the required sample size of 410 was increased to 456 to reflect an approximate 10% attrition rate. It is expected that, at most, 10% of the randomized patients will be lost-to-follow-up.

2.2 REPORTING OF RESULTS
The first CSR will be published after the analysis of Part 1 of the study is complete. This will include the results of the primary analysis and all efficacy, safety, and other data up to the Week 24 data cut-off.

The second CSR will be published after the analysis of Part 2 of the study is complete. This will include the results of all cumulative efficacy, safety, and other data up to the Week 48 data cut-off.

The final CSR will be published after the final OS analysis is complete. This analysis is scheduled to occur at the 36M/240D timepoint.

### 2.3 STUDY OBJECTIVES

#### 2.3.1 Primary Objective

**Part 1:** To compare the independently assessed best overall response rate (ORR) (according to Response Evaluation Criteria in Solid Tumor [RECIST] 1.1 criteria) at Week 24 with Mylan Trastuzumab plus taxane versus Herceptin® plus taxane in patients who have not received previous first line treatment for HER2-positive MBC.

**Part 2:** The primary objective is to descriptively compare the safety, immunogenicity, and tolerability profile of single agent Mylan Trastuzumab and Herceptin® and to compare the immunogenicity of Mylan Trastuzumab and Herceptin® by examining clinical immunogenic response.

#### 2.3.2 Secondary Objectives

**Part 1 of the study:**

- To compare independently assessed clinical activity at Week 24 between treatment arms by measuring time to tumor progression (TTP), progression-free survival (PFS), overall survival (OS), and duration of response (DR).

- To descriptively compare the safety, immunogenicity, and tolerability profile of Mylan Trastuzumab and Herceptin® given in combination with a taxane.
• To compare the populations pharmacokinetic (PopPK) AUC, $C_{\text{max}}$, minimum drug concentration ($C_{\text{min}}$), clearance, volume of distribution ($V_d$), and $T_{1/2}$ profiles of Mylan Trastuzumab and Herceptin®.

Part 2 of the study:

• To compare the clinical activity at Week 48 between treatment arms by measuring PFS, OS, and DR, and OS at 36 months or after 240 deaths from the time of randomization.

2.3.3 Exploratory Objectives

To assess the impact of shed ECD fragments of the HER2 receptor (HER2/ECD) in serum on PK and efficacy parameters.

3 TIMING OF ANALYSES AND UNBLINDING

3.1 INTERIM ANALYSIS

A formal interim analysis will be conducted when at least 30% of the information target is available. That is, the formal interim analysis will be performed after the 1st 123 randomized patients under protocol amendments 2 and beyond have either discontinued the study or completed 24 weeks in the study. The interim analyses will be comprised of just those 123 patients.

This formal interim analysis will have two objectives: (1) blinded sample size re-estimation and (2) futility analysis.

The blinded sample size re-estimation (BSSR) will take place chronologically before the interim analysis for futility, although both the BSSR and futility analysis will be based on the same group of patients. The BSSR will be conducted by a statistician independent from Mylan (e.g., the CRO blinded statistician). This person will remain blinded to the true randomization that each patient received.

The results of the BSSR may impact the cut-points used in the futility analysis. If the sample size of the study remains unchanged, then the futility analysis will be conducted as-planned by using the cut-points described in section 8.2.2. However, if the sample size of the study is increased, then Mylan (or its designee) will re-evaluate the cut-points and communicate the new cut-points to the DSMB.

The futility analysis will be unblinded and overseen by the DSMB.

A figure that depicts the conduct of the interim analysis is presented below:
3.2 TIMING OF PART 1, PART 2, AND FINAL ANALYSIS OF OS

3.2.1 Part 1: Timing and Data Cut-off
The primary efficacy analysis will be performed at the end of Part 1. If the study is not stopped after the interim analysis, the analysis of Part 1 will be performed after all randomized patients have either discontinued the study or completed 24 weeks in the study.

The data cut-off date for the analysis of Part 1 will be made at the patient level. That is, each patient will have his/her own data cut-off date, occurring at discontinuation or 24 weeks after randomization, whichever occurs first. If a study visit occurs within one week after the Week 24 data cut-off for a particular patient, that data will be included in the Part 1 analysis, and the date of that visit will become the Part 1 data cut-off date for that patient.

The analysis of Part 1 data will be limited to data points that occur on/prior or are ongoing at the time of data cut-off. Events or medications that are ongoing at the time of the data cut-off will be censored to the date of cut-off for analysis purposes.

3.2.2 Part 2: Timing and Data Cut-off
Analysis of data in Part 2 will be performed after all randomized patients have either discontinued the study or completed 48 weeks in the study.

The data cut-off date for the analysis of Part 2 will be made at the patient level. That is, each patient will have his/her own data cut-off date, occurring at discontinuation or 48
weeks after randomization, whichever occurs first. If a study visit occurs within one week after the Week 48 data cut-off for a particular patient, that data will be included in the Part 2 analysis, and the date of that visit will become the Part 2 data cut-off date for that patient.

The analysis of Part 2 data will be limited to data points that occur on/prior or are ongoing at the time of data cut-off. Events or medications that are ongoing at the time of the data cut-off will be censored to the date of cut-off for analysis purposes.

3.2.3 Timing of Final Assessment of Overall Survival (36M/240D)
Analysis of OS will be performed 36 month after the last patient was randomized into the study or after 240 patients have died, whichever occurs first.

3.3 UNBLINDING
Study unblinding will be performed at the end of Part 1 when the final analysis of the primary efficacy endpoint occurs. Only individuals fulfilling select roles at Mylan and the CRO will be unblinded (see list of roles in Blinding Plan). For both parties, blinded and unblinded teams will be established prior to the completion of study part 1. Blinded teams at Mylan and the CRO will remain blinded for the duration of the study. For Parts 1, 2 and the period of time until the final OS analysis, investigators and patients will remain blinded to the treatment that the patient received.

4 HYPOTHESES AND DECISION RULES FOR THE PRIMARY ANALYSIS

4.1 STATISTICAL HYPOTHESES
In Part 1 of the study, the ratio of the best ORRs at Week 24 will be statistically compared with the following hypotheses:

\[ H_0: \frac{R_T}{R_C} \leq 0.81 \] or \[ \frac{R_T}{R_C} \geq 1.24 \]

\[ H_1: 0.81 < \frac{R_T}{R_C} < 1.24, \]

where \( R_T \) and \( R_C \) are the best ORR of Test (Mylan Trastuzumab) and Control (Herceptin®), respectively.

4.2 STATISTICAL DECISION RULES
A two-sided 90% CI for the ratio of the best ORRs at Week 24 will be calculated based on the method of logarithmic transformation with no adjustment for covariates. The two-sided 90% CI is equivalent to two one-sided tests (TOST) at the 5% level. Equivalence will be declared if the CI is completely within the equivalence range of (0.81, 1.24). Only patients with measurable disease at baseline will be included in the analysis of the objective response.
Section 9 (DUAL SCOPE of SAP) presents the hypotheses and decision rules for EMA criteria.

5 ANALYSIS SETS

5.1 INTENT-TO-TREAT POPULATION #1 (ITT1)

The ITT Population #1 (ITT1) will consist of all patients who are randomized into the study under protocol amendment 2 and beyond. Patients in ITT1 will be categorized to the treatment as-randomized. The primary efficacy analysis will be conducted in ITT1.

5.2 INTENT-TO-TREAT POPULATION #2 (ITT2)

The ITT Population #2 (ITT2) will consist of all randomized patients. In the protocol, ITT2 is referred to as the “full ITT set”. Patients in ITT2 will be categorized to the treatment as-randomized.

5.3 PER-PROTOCOL (PP) POPULATION

The PP population will be defined at the end of Part 1 and is a subset of ITT1 who meet the following additional criteria:

- Received the treatment to which they were randomized.
- The absence of major protocol violations in Part 1 which preclude the evaluation of the patient including, for example, the lack of measurability of the lesions; the absence of violation of entry criteria which completely precludes the assessment of efficacy and safety. The details of all major protocol deviations that remove patients from PP population will be described in Blinded Data Review (BDR) Meeting Highlights.
- Has at least one post-baseline tumor assessment if a progression disease; and at least two if CR, PR or SD.
- Has received at least two complete cycles of treatment. However, if a progression, death, or discontinuation takes place because toxicity occurs before the end of the first two cycles, the patient will be retained in the PP population.

The precise reasons for excluding subjects from the PP population will be finalized at the BDR meeting to be held prior to database lock for study Part 1. Once defined for Part 1, the PP population will remain unchanged and may be used for analysis in Part 2 and for the final analysis of OS.
5.4 SAFETY (SAF) POPULATION

The SAF population will include all subjects who receive at least 1 dose of study medication (Mylan Trastuzumab or Herceptin®; in any amount), with treatment assignments designated according to actual study treatment received. The SAF population will be the primary population for the analysis of safety.

5.5 PHARMACOKINETIC (PK) POPULATION

The PopPK analysis will be based on the PK population. More detailed information can be found in the PopPK Analysis Plan.

5.6 TREATMENT MISALLOCATIONS

Misallocation of treatment will be handled in the following manner:

- If a patient is randomized but not treated, then he/she will be reported under his/her randomized treatment group for efficacy analyses. However, he/she will be excluded from the safety analyses as actual treatment is missing.
- If a patient was randomized but took incorrect treatment, then he/she will be reported under his/her randomized treatment group for all efficacy analyses but will be reported under the treatment he/she actually received for all safety analyses.

5.7 PROTOCOL DEVIATIONS

The full list of protocol deviations for the study report will be compiled prior to database closure. Timing of deviations and their relation to the statistical analyses or populations will be listed.

6 ENDPOINTS AND COVARIATES

6.1 EFFICACY ENDPOINTS

Although study conduct decisions at the patient-level will be made based on the evaluation of investigator assessments of disease response and progression, the primary and secondary analyses of endpoints dependent on disease assessments (e.g., ORR, PFS, TTP, and DR) will be based on results of the central review of disease response and progression.

Sensitivity analyses will be performed based on investigator assessments of disease response and progression.

6.1.1 Primary Endpoint
The primary efficacy endpoint is the **best overall response rate (ORR)** where objective response is defined as a CR or PR according to RECIST 1.1 based on centralized review evaluation. Objective response will be based on the best overall response recorded from Day 1 until centrally-assessed PD, death, or first administration of anti-tumor treatment (other than study medication), whichever occurs first. Patients who receive anti-tumor treatment other than the study medication prior to reaching a CR or PR, or who die, progress, or drop out for any reason prior to reaching a CR or PR will be counted as non-responders in the assessment of ORR. Patients without a post-baseline tumor assessment will be considered to be non-responders.

Only patients with measurable disease at baseline will be included in the analysis. A patient is deemed to have measurable disease if they have at least one target lesion at screening. In this protocol, bone, CNS, and skin lesions, as only metastatic disease, as well as irradiated, biopsied, or surgically manipulated lesions are excluded from being identified as measurable lesions. A tumoral skin lesion or a tumoral soft tissue lesion associated with a bone lesion could be the only measurable disease (by CT Scan or MRI) if other visceral non measurable lesions are present.

### 6.1.2 Secondary Endpoints

**Time to Tumor Progression (TTP)** is defined as the time from randomization to date of first documentation of objective progression.

Patients last known to be 1) alive; 2) on treatment or within 28 days after discontinuation of treatment; and 3) progression-free are censored at the date of the last objective disease assessment that verified lack of disease progression. Patients with no or inadequate baseline disease assessments are censored at the start date. Patients who die prior to objective progression while on treatment are censored at the date of last objective tumor assessment prior to death. Patients with documentation of progression after an unacceptably long interval (i.e., 2 or more missed or indeterminate assessments) since the last tumor assessment will be censored at the time of last objective assessment without progression.

**Progression Free Survival (PFS)** is defined as the time from randomization to first documentation of objective progression or to death due to any cause.

Patients last known to be 1) alive; 2) on treatment or within 28 days of discontinuation of treatment; and 3) progression-free are censored at the date of the last objective disease assessment that verified lack of disease progression. Patients with inadequate baseline disease assessment are censored at the start date. Patients with no on-study disease assessments are censored at the start date unless death occurred prior to the first planned assessment, in which case the death is an event. Patients with at least one on-study disease assessment who discontinue treatment without documented disease progression and without death are censored at the date of the last objective disease assessment (with objective status CR, PR or Stable). Patients with documentation of progression or death after an unacceptably long interval (2 or more missed or indeterminate assessments) since
the last tumor assessment will be censored at the time of last objective assessment without progression.

**Overall Survival (OS)** is defined as the time from date of randomization to date of death due to any cause. Patients last known to be alive are censored at date of last contact.

**Duration of Response (DR)** is defined as the time from the first documentation of objective tumor response (CR or PR) to the date of first documentation of objective tumor progression or to death due to any cause, whichever occurs first. If tumor progression data include more than 1 date, the first date will be used. Duration of tumor response will be calculated in months as (the end date for DR - first CR or PR +1)/30.44. Only patients with objective tumor response will be included in the analysis.

The censorship rules for DR are the same as that of PFS.

### 6.2 SAFETY ENDPOINTS

#### 6.2.1 Adverse Events

Assessment of adverse events will include: type, incidence, severity (graded by the National Cancer Institute [NCI] Common Terminology Criteria for Adverse Events [CTCAE], Version 4.03), timing, seriousness, and relatedness; and laboratory abnormalities. All reported AEs will be assigned the system organ class and preferred term according to Medical Dictionary for Regulatory Activities (MedDRA).

Treatment-emergent adverse events (TEAE) are defined as any AE which started or deteriorated at or after treatment with the IMP (Mylan Trastuzumab/Herceptin®) but on or within 28 days following the last dose of IMP.

#### 6.2.2 Laboratory Safety Assessments

Hematology, blood chemistry and urinalysis will be drawn at the time points described in the Schedule of Activities and analyzed at local laboratories. Investigators may have additional blood tests performed for the purpose of planning treatment administration, dose modification, or following adverse events.

Pregnancy test: Serum or urine pregnancy test for women of childbearing potential will be performed by a local laboratory.

#### 6.2.3 Other Safety Assessments

All of the following assessments will be done as outlined in the Protocol Schedule of Assessments:

**Physical examination**: A complete physical examination (excluding genitourinary examination) will be performed at selected visits. Height (at screening only) and Body weight will also be recorded.
**ECOG PS**: The ECOG scale is a 6-point categorical scale, ranging from 0 (asymptomatic) to 5 (death), and will be evaluated at selected visits.

**Vital signs**: Measurements will be made of temperature, blood pressure, heart rate, and respiratory rate.

**ECG**: A standard 12-lead ECG will be recorded at the screening visit before exposure to any of the study treatments, and then at selected time points during the study.

**Left Ventricular Ejection Fraction (LVEF)**: LVEF will be evaluated by MUGA or ECHO at the screening visit before exposure to any of the study treatment, and then at selected time points during the study.

**Immunogenicity**: Immunogenicity of Mylan Trastuzumab and Herceptin® will be assessed using validated assays in a 3-step approach:

- Step 1: Screening,
- Step 2: Confirmatory,
- Step 3: Titer,
- A validated cell-based assay will be used to assess Neutralizing Antibodies (Nab) if needed.

Samples will be taken before administration of Mylan Trastuzumab/Herceptin® since elevated antibody titer levels against Herceptin® plasma levels can interfere with the antibody assays.

### 6.3 POPULATION PHARMACOKINETIC ENDPOINTS

The analysis of the population pharmacokinetic endpoints are described in detail in the PopPK Analysis Plan.

### 6.4 COVARIATES AND STRATIFICATION FACTORS

#### 6.4.1 Stratification Factors

The randomization will be stratified by:

- Tumor progression into metastatic phase ≥2 years OR <2 years after primary diagnosis (calculated as time from primary tumor surgery until randomization). Patients diagnosed with primary metastatic disease will be classified together with the patients who progressed <2 years.
- ER/PgR status (ER and/or PgR positive/ER and PgR negative).
• Type of taxane received (e.g., paclitaxel or docetaxel). The choice of taxane will be the Investigator’s decision at the site level and will be determined before screening of the first patient at the site.

7 HANDLING OF MISSING VALUES

For the analysis of the primary endpoint, best overall response rate, if a patient qualifies for inclusion into an efficacy analysis population (ie, ITT1, ITT2, or PP) but has no post-randomization centrally-reviewed evaluation, he/she will be categorized as a non-responder.

Sensitivity analyses may include imputation of missing values. If necessary, the applicable techniques are described in section 8.6.

8 STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES

8.1 ANALYSIS SOFTWARE

SAS version 9.2 or later will be used to analyze the data from this study.

8.2 INTERIM ANALYSIS

8.2.1 Blinded Sample Size Re-estimation

A blinded sample size re-estimation will be carried out using a method due to Gould (1992) for the risk ratio (RR), and will only make use of the pooled event rate for the primary analysis of the primary endpoint. The maximum sample size that may be allowed is 600. A decrease in the required sample size via BSSR will not be permitted. Details of this methodology are as follows:

Let \( \tau_P \) denote the pooled event rate based on the patients included in this interim analysis, let \( \pi_P = 0.69 \) denote the pooled event rate as used in the sample size calculation, and let \( N = 410 \) denote the originally planned total number of randomized patients that have measurable disease at baseline. Define \( N' \) by

\[
N' = \frac{\pi_P (1 - \tau_P)}{(1 - \pi_P) \tau_P} N
\]

then the following blinded sample size re-estimation rule is defined to determine \( N^* \), the new target number of randomized patients with measurable baseline disease:

(a) If \( N' < 451 \) then \( N^* = 410 \);

(b) If \( N' > 600 \) then \( N^* = 600 \); or
(c) Otherwise $N^* = N'$ as given by equation (1).

Here, (a) indicates that there will be no sample size decrease under this procedure and no sample size change will take place if that increase would have been less than 10%. This rule corresponds to the situation when the observed pooled event rate at the interim analysis is greater than 67%. The upper limit on sample size in (b) caps the maximum sample size increase, that is, when the observed event rate is 60% or less, the sample size will be increased to a maximum of 600. When the observed ORR is within the interval of (60%, 67%), the new sample size will be calculated based on (c).

### 8.2.2 Futility Analysis

An interim analysis for futility will be conducted with the following methodology.

For the primary endpoint of best ORR at Week 24, let $\hat{\pi}_{11}$ and $\hat{\pi}_{12}$ denote the observed proportions, at this interim analysis, for the Mylan Trastuzumab and Herceptin arms, respectively, of patients in the ITT1 population with measurable disease at baseline. Then the observed log of the ratio of ORRs at this interim analysis is given by

\[ \hat{\theta}_1 = \log(\hat{\pi}_{11} / \hat{\pi}_{12}) \].

**Futility Analysis Conduct (If no increase in required sample size):**

Futility boundaries are calculated based on the timing of interim analysis and the outcome from the BSSR. For example, if the interim analysis is performed at 30% of information target, and the BSSR results in no increase of the required sample size, the trial will have met the condition to stop for futility if either $\hat{\theta}_1 < -0.303$ or $\hat{\theta}_1 > 0.303$ based on this interim data. That is, if the observed ratio of the ORR at the interim analysis is found to be less than 0.739 or greater than 1.354, the trial will have met the condition to stop for futility. This boundary has been derived so that the decrease in power of the final analysis due to introducing the interim analysis for futility is negligible (at most 1%). The table below gives the probability of stopping for futility for a wide range of true values of ratio of ORR and log RR when the true value of the pooled event rate is 69%.

<table>
<thead>
<tr>
<th>True Ratio of ORR</th>
<th>True Log RR</th>
<th>Percentage of Trials that would Stop at Interim for Futility</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.779</td>
<td>-0.25</td>
<td>39.3%</td>
</tr>
<tr>
<td>0.819</td>
<td>-0.20</td>
<td>26.0%</td>
</tr>
<tr>
<td>0.861</td>
<td>-0.15</td>
<td>15.4%</td>
</tr>
<tr>
<td>0.905</td>
<td>-0.10</td>
<td>8.2%</td>
</tr>
<tr>
<td>0.951</td>
<td>-0.05</td>
<td>4.3%</td>
</tr>
<tr>
<td>1.000</td>
<td>0.00</td>
<td>3.1%</td>
</tr>
</tbody>
</table>
Note: When the true ratio of ORR is 1, only 1% of such trials would stop for futility in cases where equivalence would have been obtained at the end of the study, and in the other 2.1% of cases, equivalence would not have been demonstrated at the end of the study.

**Futility Analysis Conduct (If required sample size is increased):**
If the BSSR results in an increase in the sample size, then Mylan (or its designee) will re-evaluate and communicate the new cut-points for \( \hat{\Theta} \) to the DSMB prior to its conduct of the futility analysis. The communication will be in the form of a memo and will include the newly defined futility cut-points and the above table with newly calculated stopping probabilities associated with various log RR.

### 8.3 DESCRIPTIVE SUMMARIES

Unless otherwise specified, descriptive data summaries will be tabulated by treatment for all endpoints. Categorical outcomes will be summarized by number and percent of patients that fall into each category. For continuous outcomes, descriptive statistics including number of subjects (n), arithmetic mean, standard deviation, minimum, median, maximum, and 95% CIs of the means will be presented.

Where data are collected over time, both the observed data and the change from baseline will be summarized at each visit.

### 8.4 PRIMARY ANALYSIS

The primary efficacy endpoint is the best ORR where objective response is defined as a CR or PR according to RECIST 1.1 based on central tumor evaluation. The ratio of the best ORRs at Week 24 will be statistically compared with the following hypotheses:

\[
H_0: \left( \frac{R_T}{R_C} \leq 0.81 \right) \text{ or } \left( \frac{R_T}{R_C} \geq 1.24 \right)
\]

\[
H_1: 0.81 < \left( \frac{R_T}{R_C} \right) < 1.24,
\]

where \( R_T \) and \( R_C \) are the best ORR of Test (Mylan Trastuzumab) and Control (Herceptin®), respectively.

A two-sided 90% CI for the ratio of the best ORRs at Week 24 will be calculated based on the method of logarithmic transformation with no adjustment for covariates. The two-

---

This Document is Company Confidential.
sided 90% CI is equivalent to two one-sided tests (TOST) at the 5% level. Equivalence will be declared if the CI is completely within the equivalence range of (0.81, 1.24). A two-sided 95% CI will also be calculated for exploratory purpose.

In this protocol, bone, CNS, and skin lesions, as only metastatic disease, as well as irradiated, biopsied, or surgically manipulated lesions are excluded from being identified as measurable lesions. A tumoral skin lesion or a tumoral soft tissue lesion associated with a bone lesion could be the only measurable disease (by CT Scan or MRI) if other visceral non measurable lesions are present.

8.5 SECONDARY ANALYSES

The following secondary analyses will be performed at Week 24 and 48 for Part 1 and Part 2, respectively, for the ITT1 population:

- For TTP, PFS, OS, and DR, Kaplan-Meier plots by treatment will be presented and the log-rank test of the two treatment groups unadjusted for any covariates will be performed.
- For TTP, PFS, OS, and DR, Cox’s proportional hazards model will be used to analyze for treatment effects, adjusting for subgroup. Univariate analysis and multivariate analysis with forward selection will be performed. Hazard ratios and 95% CIs will be presented. Forest plots will be produced for subgroups. The subgroups are identified as factors where the efficacy results may differ from category to category within the subgroup. The subgroups are stratification factors, age, race, previous adjuvant/neoadjuvant chemotherapy or HER2 targeted treatment, visceral metastases, number of metastatic sites, CNS as first site of metastasis and geographic region (North America, Western Europe, Eastern Europe, Asia Pacific, Africa, Latin America).

Additionally, OS will be evaluated at the 36M/240D timepoint in the same manner.

DR will be analyzed in the same manner as TTP at Week 48 only. Only patients with objective response (PR or CR) will be included in the analysis.

8.6 SENSITIVITY ANALYSES

8.6.1 Sensitivity Analyses for ORR

The following sensitivity analyses will be performed on the primary endpoint, ORR:

- Subgroup analyses by stratification factors, age, race, previous adjuvant/neoadjuvant chemotherapy or HER2 targeted treatment, visceral metastases, number of metastatic sites, CNS as first site of metastasis and geographic region (North America, Western Europe, Eastern Europe, Asia Pacific, Africa, Latin America). Forest plots will be produced for subgroups.
• Univariate analysis and multivariate logistic regression analysis with forward selection will be performed based on the factors and covariates indicated above.

• Cochran-Mantel-Haenszel analysis stratified by the stratification factors will be performed. Estimates of the relative risk and the odds ratio, and their 90% and 95% CIs will be presented.

• Logistic regression analysis of the treatment odds ratio adjusted for the stratification factors will be performed.

• The primary efficacy analysis will be replicated in the PP population.

• The primary efficacy analysis will be replicated in the ITT2 population.

• The primary efficacy analysis will be conducted with the investigator assessments of disease response and progression in the ITT1 population.

• The difference in ORR (with a two-sided 90% and 95% CIs) will be calculated with no covariate adjustment in the ITT1 population. Equivalence will be evaluated within the equivalence region of (-15%, +15%).

8.6.2 Sensitivity Analyses for TTP

• All secondary analyses for TTP will be replicated in the ITT2 population.

• All secondary analyses for TTP will be replicated in the PP population.

• All secondary analyses for TTP will be replicated in the ITT1 population with investigator assessments of disease response and progression.

8.6.3 Sensitivity Analyses for PFS

• All secondary analyses for PFS will be replicated in the ITT2 population.

• All secondary analyses for PFS will be replicated in the PP population.

• All secondary analyses for PFS will be replicated in the ITT1 population with investigator assessments of disease response and progression.

• Given that this study is ongoing and survival data continues to be collected, preliminary evaluation of PFS based on investigator assessment will also be presented.

8.6.4 Sensitivity Analyses for OS

• All secondary analyses for OS will be replicated in the ITT2 population.

• All secondary analyses for OS will be replicated in the PP population.
• A truncated OS (tOS) analysis will be conducted at Year 2 in the ITT1 population. Any patient still alive after two years will be censored at two years. Any patient who died after two years will be censored at two years.

8.7 EXPLORATORY ANALYSIS

8.7.1 Disease Control Rate

Disease control rate (CR, PR, or SD) at Week 24 and 48 will be presented by treatment group. The ratio of the rate and its 90% and 95% CIs will also be presented.

8.7.2 HER2/ECD

HER2/ECD will be analysed in the following manner:

**Evaluating Baseline HER2/ECD as a Predictor:** Baseline HER2/ECD will be evaluated as predictor for the efficacy endpoints of ORR, OS, and TTP at Week 24 and 48.

- One-Way Model: For ORR, a one-way logistic regression model will be built at Week 24 and Week 48, and the p-value for baseline HER2/ECD will be presented for each timepoint. For OS and TTP, one-way Cox PH models will be built at Week 24 and 48, and the p-value for baseline HER2/ECD will be presented for each timepoint.
  - Interaction Model: Interaction models with treatment, baseline HER2/ECD, and treatment-by-baseline HER2/ECD interaction will be built for ORR, OS, and TTP at Week 24 and 48. The p-value for each factor in the model will be presented.

**General Descriptive Summaries of HER2/ECD:** The following descriptive summaries will be presented:

- Summary of HER2/ECD by visit and treatment
- Summary of change from baseline in HER2/ECD by visit and treatment
- Summary of percent change from baseline in HER2/ECD by visit and treatment

**Elevated HER2/ECD:** Elevated HER2/ECD will be defined as a HER2/ECD value of 15 ng/ml or greater and will be descriptively summarized by visit and treatment group. Rates of ORR will be descriptively summarized by baseline HER2/ECD elevation status and by baseline HER2/ECD elevation status and treatment.

**Significant Percent Change from baseline in HER2/ECD:** A significant percent change from baseline in HER2/ECD will be defined as a decrease of 55% or more. Significant percent change status will be descriptively summarized by visit and treatment group. Rates of ORR will be descriptively summarized by significant percent change status at each visit and by significant percent change status at each visit by treatment.
8.8 ANALYSIS OF OTHER ENDPOINTS

Descriptive statistics will be used to summarize subject disposition, baseline characteristics, and treatment administration/compliance.

**Subject Disposition** - The number of subjects screened, enrolled, treated, evaluated for safety, the number of subjects who completed the study, the number of subjects who failed the screen and reasons for screen failures, and the number of subjects who withdrew from the study and reasons for discontinuation will be summarized by treatment arm.

A KM plot of the time from randomization to discontinuation from the study by treatment will be presented at the end of Part 1, Part 2, and final OS analysis.

**Demographic and Baseline Characteristics** - Demographic and baseline characteristics such as age, gender, height, weight, race, prior therapy, prior medication, physical examination, ECOG performance status, signs and symptoms, geographical region, and medical/oncology history will be tabulated by treatment.

**Disease History** – The following characteristics will be summarized by treatment group: visceral metastases, number of metastatic sites, and brain as first site of metastasis.

**Treatment Administration/Compliance** - Study drug administration will be summarized for each treatment at the end of Part 1 (Mylan Trastuzumab, Herceptin®, and taxane), Part 2 (Mylan Trastuzumab and Herceptin®). The number of subjects with interruptions in dosing and number of subjects with reduction in dosing and reasons for the deviations from planned therapy will be summarized by treatment.

8.9 SAFETY ANALYSES

All safety analyses will be performed on the SAF population as-treated. All subjects who receive any study medication will be included in the summaries and listings of safety data. Summaries of TEAEs and other safety parameters will be provided as appropriate.

Safety analyses for Part 2 will be carried out on cumulative data of all patients from Part 1. In addition, a separate analysis will be done on the patients who entered Part 2 of the study and took monotherapy only during Part 2. For those patients, the analyses will be performed using their Part 2 data only.

8.9.1 Adverse Events

All reported AEs will be assigned the system organ class and preferred term according to Medical Dictionary for Regulatory Activities (MedDRA), and graded by CTCAE Version 4.03. Listing of all AEs will be tabulated by treatment groups and by system organ class and preferred term.
Treatment-emergent adverse events (TEAE) are defined as any AE which started or deteriorated at or after treatment with the IMP (Mylan Trastuzumab/Herceptin®) but on or within 28 days following the last dose of IMP.

Patient incidence of the following AEs will be tabulated by treatment groups and by system organ class and preferred term:

- All TEAEs by worst grade.
- Grade 3 or higher TEAEs.
- SAEs.
- TEAEs leading to discontinuation of IMP.
- TEAEs leading to interruption of IMP.
- TEAEs leading to removal from the study.
- Fatal AEs.
- All above analyses will be replicated for treatment related AEs.

In addition, the exposure-adjusted incidence rates of selected categories of AEs may be tabulated by treatment group.

8.9.2 Clinical Laboratory Data

For hematology and biochemistry variables, descriptive summaries of observed values and changes from baseline will be presented by treatment arm.

Each abnormal value will be flagged to show whether it is a value below or above the reference range.

The assessments of laboratory variables will be tabulated by visit for each clinical laboratory parameter by treatment arm (frequency tables). Additionally, for each laboratory parameter, shifts in assessments from baseline to all post-baseline visits will be presented by treatment arm (shift tables).

Laboratory values that are outside the reference range will also be flagged in the data listings, along with corresponding reference ranges.

The assessment of categorical urinalysis variables will be tabulated by visit for each urine parameter by treatment arm (frequency tables). Additionally, for each of these urine parameter shifts in assessments from baseline to all post-baseline visits will be presented for each treatment arm (shift tables).

8.9.3 Vital Signs

Descriptive summaries of observed values and changes from baseline will be calculated for vital signs. These summaries will be presented by visit and treatment arm.

This Document is Company Confidential.
8.9.4 Electrocardiograms
Descriptive summaries of observed values and changes from baseline will be calculated for ECG variables. QT will be corrected according to both Fridericia and Bazett formulas. Frequency and shift tables will be presented for the classified values of QTc as given by International Conference on Harmonization (ICH) E14 as well as for the overall clinical assessment. The ICH-E14 classifications are as follows:

- Absolute QTc interval prolongation:
  - QTc interval >450.
  - QTc interval >480.
  - QTc interval >500.
- Change from baseline in QTc interval:
  - QTc interval increases from baseline >30.
  - QTc interval increases from baseline >60.

8.9.5 Left Ventricular Ejection Fraction
Descriptive summaries of observed values and changes from baseline will be calculated for LVEF. These summaries will be presented by visit and treatment arm.

8.10 ECOG PS
ECOG performance status and change from baseline in ECOG PS will be summarized at each scheduled visit where it is collected.

Change from baseline in ECOG PS will be summarized as a continuous variable and as a categorical variable. A decrease of one point or more from baseline will be categorized as an “improvement” from baseline. An increase of one point or more from baseline will be categorized as a “deterioration” from baseline. Improvement, Deterioration, and Unchanged ECOG PS from baseline will be summarized as a categorical variable by treatment at each post-randomization timepoint that ECOG PS is evaluated.

8.11 PK
The PopPK analyses described in the protocol will be detailed in the PopPK Analysis Plan.

8.12 IMMUNOGENICITY
Immunogenicity is characterized by patients who test positive for Human Anti-Human Antibodies (HAHA) against trastuzumab or Mylan trastuzumab.

A descriptive summary of the number and percentage of patients who test positive for HAHA by treatment will be produced.
For each patient who tests positive for HAHA, the safety listings will be reviewed for related AEs:

If the total number of ADA-positive patients is 30 or more, then a descriptive summary of the number of patients (and percentage) who experience the above safety-events-of-interest by treatment and HAHA status will be produced. The demographic and baseline characteristics will also be summarized for these patients.

8.13 MULTIPLICITY ADJUSTMENTS

No multiple comparison adjustment for the Part 1 primary analysis is required. All other efficacy analyses for Part 1 or 2 will not be adjusted for multiplicity.

8.14 THE INTERPRETATION OF PART 2 ANALYSIS RESULTS

Part 2 data may be analyzed separately from Part 1 or cumulatively across Parts 1 and 2. In either case, careful consideration will be made when interpreting the results from these analyses. If the Part 2 analyses are affected, biased, or weakened due to attrition, barriers to entry into Part 2, or other factors, then these concerns will be communicated in the CSR upon the presentation of these analysis results. In a worst case scenario, some Part 2 analyses will be considered descriptive.

9 DUAL SCOPE OF SAP

This SAP is primarily written for the US regulatory authority (US FDA), which recommended the ratio of best ORRs as the primary efficacy analysis. In contrast, the European regulatory authority (EMA) has requested that the difference in best ORRs is used as the primary efficacy analysis. The current sample size was chosen to adequately power each primary efficacy analysis. Section 2.1.1 details the sample size calculations for both the US FDA- and EMA-recommended primary efficacy analyses.

The hypotheses and decision rule for the US FDA-recommended primary efficacy analysis of the ratio of best ORRs are presented in Sections 4.1 and 4.2. The analogous hypotheses and decision rule for the EMA-recommended primary efficacy analysis of the difference in best ORRs are as follows:

With the EMA’s equivalence margin of (-15%, 15%), the hypotheses are set up as following:

H₀: (Rₜ - Rₖ ≤ -15%) or (Rₜ - Rₖ ≥ 15%)

H₁: -15% < (Rₜ - Rₖ) < 15%,

where Rₜ and Rₖ are the best ORR of Test (Mylan Trastuzumab) and Control (Herceptin®), respectively. A two-sided 95% CI for the difference of the best ORRs at
Week 24 will be calculated. Equivalence can be declared if the CI is completely within the equivalence range of (-15%, 15%).

10 REFERENCES


FDA (February 2010), Draft Guidance for Industry, "Adaptive Design Clinical Trials for Drugs and Biologicals"

Gould A.L. (1992), "Interim analyses for monitoring clinical trials that do not materially affect the Type I error Rate", Statistics in Medicine, 11: 55-66

Appendix A

Estimation of the Treatment Effect by Fixed-Effects Meta-Analysis of Historical Herceptin Studies and Justification of the Equivalence Region for MYL-HER 3001

1. Selected studies and populations:

Fifteen relevant trials were reviewed (Prof. Dr. Lehert [1]), from which 3 trials proved to be appropriate as randomized trials comparing taxane (TX) and trastuzumab (TZ) combination versus taxane alone. To be consistent with the current study population, only the IHC3+ and/or FISH positive patient populations were included in the conducted meta-analysis. The selected three studies are:

a. Gasparini 2007 - Randomized, Phase II Trial of weekly paclitaxel alone versus trastuzumab plus weekly paclitaxel as first-line therapy of patients with Her-2 positive advanced breast cancer; open label, randomized study; population: aged >18, ICH2+ or 3+, ECOG ≤2, previous anthracycline, taxane 12 months prior to study start allowed, patients have received chemotherapy for metastatic disease excluded.

b. Slamon 2001 - Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. Population: ICH2+ or 3+ women, previous anthracycline allowed, previous chemotherapy for metastatic disease not allowed.

c. Marty 2005 - Randomized Phase II trial of the efficacy and safety of trastuzumab combined with docetaxel in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer administered at first-line treatment; open label, randomized study; population: aged 18-70, patients with any prior treatment of metastatic disease with taxanes or anti-HER therapy excluded, ECOG ≤2, ICH2+ or 3+, amended to include 3+ only.

2. Meta-analysis of IHC3+ and/or FISH-positive patients from the three randomized studies

The p-value for testing the heterogeneity among the three selected studies was 0.297, indicating that the heterogeneity among the three studies in the meta-analysis was not significant. Hence, a fixed-effects approach was used for the meta-analysis which included only the IHC3+ and/or FISH positive patient populations in the selected trials. Table 1 presents the meta-analysis results.
### Table 1: IHC3+ and/or Fish Positive Patients from 3 Randomized Trials

<table>
<thead>
<tr>
<th>Study</th>
<th>TX+TZ Events</th>
<th>TX+TZ Subjects</th>
<th>TZ Events</th>
<th>TZ Subjects</th>
<th>95% CI for Risk Ratio (RR&lt;sup&gt;1&lt;/sup&gt;)</th>
<th>Lower</th>
<th>Upper</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gasparini</td>
<td>33</td>
<td>39</td>
<td>20</td>
<td>42</td>
<td>1.777</td>
<td>1.259</td>
<td>2.507</td>
<td>0.3991</td>
</tr>
<tr>
<td>Marty</td>
<td>56</td>
<td>92</td>
<td>32</td>
<td>94</td>
<td>1.788</td>
<td>1.291</td>
<td>2.476</td>
<td>0.4461</td>
</tr>
<tr>
<td>Slamon</td>
<td>33</td>
<td>68</td>
<td>13</td>
<td>77</td>
<td>2.874</td>
<td>1.654</td>
<td>4.996</td>
<td>0.1548</td>
</tr>
<tr>
<td>Overall</td>
<td>Estimate of RR&lt;sup&gt;1&lt;/sup&gt;</td>
<td>1.92</td>
<td>95% CI for Overall RR&lt;sup&gt;1&lt;/sup&gt;</td>
<td>(1.544 , 2.386)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: 1. RR<sup>1</sup> = ORR of (TX+TZ) divided by ORR of TX. 2. SAS software 9.3 was used for the above calculation.

### 3. Equivalence margin selection

The results from the meta-analysis were used to statistically justify the equivalence margin in study MYL-HER3001. From the meta-analysis, the estimated treatment effect and its 95% CI was 1.92 (1.544, 2.386). A retention of 50% of the treatment effect is a typical basis for an equivalence margin.

Thus, if the lower bound of the 95% CI (ie, 1.544) is used to conservatively estimate the treatment effect, then a retention of 50% of that effect is 1.24*. This value will be the upper bound of the equivalence region in MYL-HER3001. The lower bound of the equivalence region is 0.81 (ie, 1/1.24=0.81). The equivalence region (0.81, 1.24) is symmetrical on the natural log scale.

* The value 1.24 is derived as follows. Firstly, 1.544 is transformed to the natural log scale. This value is 0.43438. Half of that value is 0.21719. This equates to a 50% retention of the effect size. Exponentiating that value gives 1.24.

### References


Signature Page for Statistical Analysis Plan (SAP)

Approval of SAP—Version 2.2

Author:

Jinyu Yuan, Ph.D.
Director, Biostatistics

Approval:

Ashwani Marwah
Associate Scientific Manager, Biostatistics
Biocon Research Limited

Date

Eduardo Pennella, MD, MBA
Senior Director, Oncology/Hematology
Mylan Inc

Date

Hans-Friedrich Koch, Ph.D.
Senior Director, Biometrics
Mylan Inc

Date

This Document is Company Confidential.
Signature Page for Statistical Analysis Plan (SAP)

Approval of SAP – Version 2.2

Author:

Jinyu Yuan, Ph.D.
Director, Biostatistics

Approval:

Ashwani Marwah
Associate Scientific Manager, Biostatistics
Biocon Research Limited

Eduardo Pennella, MD, MBA
Senior Director, Oncology/Hematology
Mylan Inc

Hans-Friedrich Koch, Ph.D.
Director, Biometrics
Mylan Inc

Date

Date

Date

22-JUN-2016

This Document is Company Confidential.
Signature Page for Statistical Analysis Plan (SAP)

Approval of SAP – Version 2.2

Author:

Jinyu Yuan, Ph.D.
Director, Biostatistics

Approval:

Ashwani Marwah
Associate Scientific Manager, Biostatistics
Biocon Research Limited

Eduardo Pennella, MD, MBA
Senior Director, Oncology/Hematology
Mylan Inc

Hans-Friedrich Koch, Ph.D.
Director, Biometrics
Mylan Inc

This Document is Company Confidential.
Signature Page for Statistical Analysis Plan (SAP)

Approval of SAP – Version 2.2

Author:

Jinyu Yuan, Ph.D.
Director, Biostatistics

Date

Approval:

Ashwani Marwah
Scientific Manager, Biostatistics
Biocon Research Limited

Date: 25.08.2016

Eduardo Pennella, MD, MBA
Senior Director, Oncology/Hematology
Mylan Inc

Date

Hans-Friedrich Koch, Ph.D.
Director, Biometrics
Mylan Inc

Date

MYL-Her 3001 Statistical Analysis Plan Version 2.2

This Document is Company Confidential.