NOVARTIS

Clinical Development

LAP016/Lapatinib/Tykerb®

CLAP016A2206 / NCT02213042

An Open-Label, Phase II, Study to Evaluate Biomarkers Associated with Response to Subsequent Therapies in Subjects with HER2-Positive Metastatic Breast Cancer Receiving Treatment with Trastuzumab in Combination with Lapatinib or Chemotherapy (EGF117165)

Statistical Analysis Plan (SAP) for final CSR

Author: Statistician: [redacted]

Document type: SAP Documentation

Document status: Final Amendment 2

Release date: 10-NOV-2020

Number of pages: 32

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
10-JUN-2020	Prior to DB lock	Creation of final version	N/A - First version for final analysis	NA
24-JUL-2020	Prior to DB lock	Creation of Amendment 1	Revised MedDRA version # The study team planned to upgrade MedDRA to newer version. However, there was a concern that implementing MedDRA upgrade might invalidate investigator signatures for certain closed sites. Consequently the team agreed to not upgrade MedDRA.	Section 2.8.1
10-NOV-2020	Prior to DB lock	Creation of Amendment 2	Revised the text for producing the outputs for regulatory disclosure	Section 2.8.1 Section 2.8.3

Table of contents

Table of contents 3
List of abbreviations 5
1 Introduction 7
1.1 Study design 7
1.2 Study objectives and endpoints 9
1.2.1 Primary objective 9
1.2.2 Secondary objectives 9
[redacted] [redacted] 10
1.2.4 Primary endpoint 10
1.2.5 Secondary endpoints 10
[redacted] [redacted] 10
[redacted] [redacted] 10
2 Statistical methods 10
2.1 Data analysis general information 10
2.1.1 Data analysis 10
2.1.2 General definitions 11
2.2 Analysis sets 14
2.2.1 Subgroup of interest 15
2.3 Patient disposition, demographics and other baseline characteristics 15
2.3.1 Patient disposition 15
2.3.2 Protocol deviations 15
2.3.3 Demographic and baseline characteristics 15
2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance) 15
2.4.1 Study treatment / compliance 15
2.4.2 Extent of exposure 16
2.4.3 Prior, concomitant and post therapies 17
2.5 Analysis of the primary objective 17
2.5.1 Primary endpoint 17
2.5.2 Statistical hypothesis, model, and method of analysis 17
2.5.3 Handling of missing values/censoring/discontinuations 19
2.5.4 Supportive analyses 19
2.6 Analysis of the secondary objective 19
2.6.1 Secondary endpoint 19
[redacted] [redacted] 23
2.8 Safety analyses 23

2.8.1 Adverse events (AEs) 23
2.8.2 Adverse events of special interest / grouping of AEs 25
2.8.3 Deaths and serious adverse events (SAEs) 27
2.8.4 Adverse events leading to discontinuation of study treatment and/or withdrawal from the study and other significant adverse events 27
2.8.5 Clinical laboratory evaluations 28
2.8.6 Analyses of liver function test 28
2.8.7 Liver toxicity 29
2.8.8 Other safety measures 29
2.9 Pharmacokinetic endpoints 31
2.10 Patient-reported outcomes 31
2.11 Biomarkers 31
2.12 Interim analysis 32
3 Sample size calculation 32
4 Change to protocol specified analyses 32
5 Reference 32

List of abbreviations

AE	Adverse Event
AI	Aromatase Inhibitor
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Classification
CBR	Clinical Benefit Rate
CI	Confidence Interval
CR	Complete Response
CSR	Clinical Study Report
CT	Computed Tomography
CTCAE	Common Terminology Criteria for Adverse Events
DBL	Database Lock
DILI	Drug Induced Liver Injury
ECG	Electrocardiogram
ECHO	Echocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EMA	European Medicines Agency
EOS	End of Study
EP	Evaluable Population
ER	Estrogen Receptor
EU	Europian Union
FAS	Full Analysis set
GCP	Good Clinical Practice
HER2	Human Epidermal growth factor Receptor 2
IHC	Immunohistochemistry
IP	Investigational Product
ITT	Intent-to Treat
IV	Intra-Venous
MBC	Metastatic Breast Cancer
MedDRA	Medical Dictionary for Drug Regulatory Affairs
MRI	Magnetic Resonance Image
NCI	National Cancer Institute
NE	Not Evaluable
ORR	Overall Response Rate
PAM50	Prediction Analysis of Microarray 50
PD	Progressive Disease
PFS	Progression-Free Survival
PP	Per-Protocol
PR	Partial Response
RECIST	Response Evaluation Criteria in Solid Tumors

RUCAM	Roussel Uclaf Causality Assessment Method
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SD	Stable Disease
SOC	System Organ Class
TIL	Tumor Infiltrating Lymphocytes
LVEF	Left Ventricular Ejection Fraction

1 Introduction

This statistical analysis plan (SAP) is for the final clinical study report (CSR). It is based on the SAP Amendment 1 for the primary CSR of CLAP016A2206 (originally the GSK study EGF117165) with minimal updates. CLAP016A2206 is a multicenter, open-label, Phase II study in patients with HER2-positive metastatic breast cancer who received at least 2 prior lines of anti-HER2-targeted therapies of which at least one included a trastuzumab-containing regimen. The primary objective is to evaluate the changes in the expression of biomarkers associated with immunomodulation between the pre-treatment biopsy and the progression biopsy and within each arm. Secondary efficacy objectives are to evaluate overall response rate (ORR); clinical benefit rate (CBR); and progression-free survival (PFS) on study treatment; to explore association between changes in biomarkers and PFS; as well as safety/tolerability. The recruitment was stopped on 31 March 2017due to difficulty in enrolling patients, and the study was terminated early. There were 42 of the 225 planned subjects enrolled. Primary analysis had been performed with data cut-off date of 27-Mar-2018 and treatment was ongoing for 4 subjects at the time. In this SAP, references will be made to the following documents:

- Protocol Amendment 03, 2013N170247_03 (effective 15May2017)
- Electronic Case Report Form (eCRF) for EGF117165
- Statistical Analysis Plan Amendment 1 for EGF117165 primary CSR

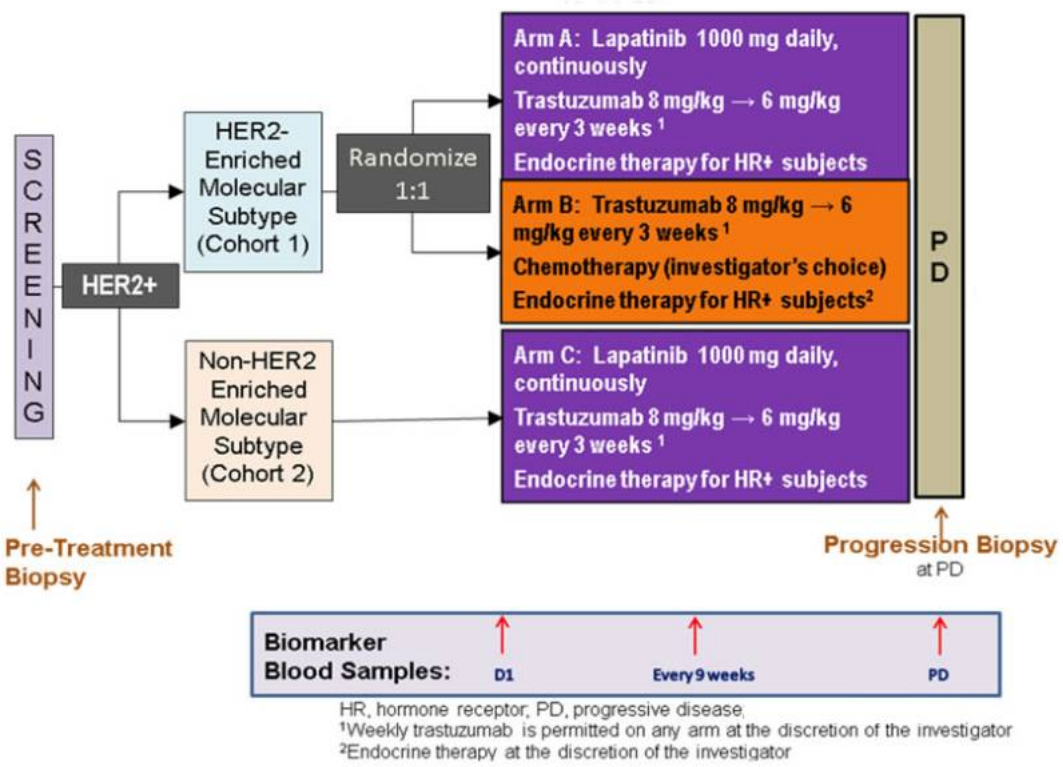
1.1 Study design

This is a Phase II, 3 arm, open-label study (see Figure 1-1 for study schema). The study population is comprised of patients with HER2-positive metastatic breast cancer who received at least two prior regimens containing trastuzumab in the metastatic setting. The last regimen must have included trastuzumab plus chemotherapy. The molecular subtype of a biopsy of a metastatic site will be centrally determined by the Prosigna assay. HER2 and hormone receptor status will be centrally determined. Patients with centrally determined HER2 positive disease will be allocated to one of two cohorts depending on the molecular subtype of their biopsy:

- COHORT 1 - HER2-Enriched contains HER2+ patients with a HER2-enriched molecular subtype. Patients with HER2-enriched molecular subtype will be randomized 1:1 to either:
  - ARM A - Trastuzumab in combination with lapatinib: lapatinib 1000 mg PO once daily plus trastuzumab (loading dose of 8 mg/kg) followed by the maintenance dose of 6 mg/kg IV q3weekly. Weekly trastuzumab (loading dose of 4 mg/kg) followed by maintenance dose of 2 mg/kg IV weekly is acceptable at the discretion of the investigator. Patients with hormone receptor (ER and/or PgR)-positive, HER2-positive MBC (metastatic breast cancer) randomized to this arm will be required to be treated with an aromatase inhibitor of the investigator's choice.
  - ARM B - Trastuzumab in combination with chemotherapy: trastuzumab (loading dose of 8 mg/kg) followed by maintenance dose of 6 mg/kg IV q3weekly plus chemotherapy of the investigator's choice. Weekly trastuzumab (loading dose of 4 mg/kg) followed by maintenance dose of 2 mg/kg IV weekly is acceptable at the discretion of the investigator. An aromatase inhibitor may be used for hormone receptor-positive, HER2-positive patients at the discretion of the investigator.

- COHORT 2 - Non-HER2-Enriched contains HER2+ patients with luminal A, luminal B and basal-like molecular subtypes
  - Patients in the Non-HER2 enriched cohort will be assigned to a third arm: ARM C - Trastuzumab in combination with lapatinib: lapatinib 1000 mg PO once daily plus trastuzumab (loading dose of 8 mg/kg) followed by the maintenance dose of 6 mg/kg IV q3weeks. Weekly trastuzumab (loading dose of 4 mg/kg) followed by maintenance dose of 2 mg/kg IV weekly is acceptable at the discretion of the investigator. Patients with hormone receptor (ER and/or PgR)-positive, HER2-positive MBC randomized to this arm will be required to be treated with an aromatase inhibitor of the investigator's choice.

Figure 1-1 Study Schema



Biopsies of a metastatic site will be collected at two time points: screening (pre-treatment) and at disease progression (PD). The primary objective is to evaluate changes in biomarkers associated with immunomodulation, between the pre-treatment biopsy and the progression biopsy within each arm. Radiologic disease assessments will be performed every 9 weeks until 54 weeks then every 24 weeks thereafter until disease progression, death, or withdrawal from study treatment for any reason (e.g. unacceptable toxicity). [redacted]

Safety assessments including physical examination, Eastern Cooperative Oncology Group (ECOG) performance status, vital signs and weight, adverse event (AE) monitoring, and

laboratory tests (complete blood count, blood chemistry including liver function test) will be done at screening and then every 3 weeks for the duration of therapy. Additional safety assessments, such as cardiac monitoring, are required for all patients every 12 weeks until discontinuation from the study therapy. All patients will receive study treatment until disease progression, death, unacceptable toxicity, and withdrawal of consent or any other reasons.

Study completion

A subject will be considered to have completed the study if the subject presents with disease progression, starts a new anti-cancer therapy, dies or withdraws from the study, or the study ends, whichever comes first. In case of disease progression during the treatment period, the subject will be followed-up for 30 days for safety evaluation and no further study-specific follow-up will be carried out; the subject will be considered as having completed the study.

In case of study treatment discontinuation for any reason other than disease progression, the subject will be followed-up for safety and efficacy assessments until disease progression, new anticancer therapy, death, withdrawal of consent or end of study, whichever comes first.

Following the data cut-off date for the primary analysis on 27-Mar-2018, the study remained open. Ongoing patients continued to receive study treatment and be followed as per the schedule of assessments, as long as patients derive benefit from lapatinib. The end of study is defined as the earliest occurrence of one of the following:

- All patients have died or discontinued from the study
- Another clinical study becomes available that can continue to provide lapatinib in this patient population and all patients ongoing are eligible to be transferred to that clinical study

1.2 Study objectives and endpoints

1.2.1 Primary objective

To evaluate changes in the expression of biomarkers associated with immunomodulation.

1.2.2 Secondary objectives

- To describe overall response rate (ORR), clinical benefit rate (CBR), and progression free survival (PFS) on study treatment in patients treated with trastuzumab in combination with lapatinib or chemotherapy.
- To explore association between changes in biomarkers and PFS.
- To describe the safety and tolerability of trastuzumab in combination with lapatinib and of trastuzumab in combination with chemotherapy.

Analysis for correlation between changes in expression profile of biomarkers from baseline to disease progression and PFS will not be conducted due to insufficient sample size from the early closure.

[redacted]

1.2.4 Primary endpoint

The primary efficacy endpoint of this study is the change in expression profile of genes and/or proteins from baseline to disease progression biopsies within each arm on a set of biomarkers associated with immunomodulation. As described above, the HER2-enriched cohort (Arm A and Arm B) will contain HER2+ patients with a HER2-enriched molecular subtype determined by PAM50; while HER2-non enriched cohort (Arm C) will contain HER2+ patients with the luminal A, luminal B and basal-like molecular subtypes.

1.2.5 Secondary endpoints

The secondary endpoints of this study are:

- Investigator-assessed PFS (defined as the interval of time between randomization and disease progression or death due to any cause);
- Investigator-assessed ORR (defined as percentage of patients with a CR or PR);
- CBR (defined as percentage of patients with a CR, PR, or SD for at least 6 months);
- AE, laboratory parameters, ECG and vital signs

[redacted]

2 Statistical methods

2.1 Data analysis general information

The final analysis will take place once all patients have discontinued treatment and completed the study. All the data collected during the entire period of the study will be included in the final analysis.

2.1.1 Data analysis

- All analyses and outputs will be performed and produced by Novartis;
- All programming will be performed using SAS version 9.4 or a later release, in a UNIX environment;
- The treatments were assigned to patients using a central randomization independent of centers. As it is anticipated that accrual will be spread thinly across centres and

summaries of data by centre would be unlikely to be informative, data from all participating centres will be pooled prior to analysis.

- Continuous variables will be summarized with the statistics mean, median, standard deviation (SD), minimum and maximum, and categorical variables will be summarized with frequency counts and percentages;
- All confidence intervals will be two-sided and will use 95% confidence levels. If sample size permits, any analysis requiring significance testing will use a two-sided test at the 0.05 significance level, unless otherwise specified;
- P-values will be rounded to three decimal places. P-values that round to 0.000 will be presented as '<0.001' and p-values that round to 1.000 will be presented as '>0.999'. Any p-value considered statistically significant will be marked with one asterisk (e.g., 0.022*);
- All listings will be sorted for presentation in order of treatment regimen, country, study center, patient, and date of procedure or event;
- All analysis and summary tables will have the analysis population sample size (i.e., number of patients);
- Visits outside the protocol defined time-windows (i.e. recorded as protocol deviations) will be included in listings but omitted from figures, summaries and statistical analyses.

2.1.2 General definitions

2.1.2.1 Treatment groups

Table 2-1 Treatment groups

Treatment Arm	Descriptor	Table Order & Label
COHORT 1 - HER2-Enriched contains HER2+ patients with a HER2-enriched molecular subtype. Patients with HER2-enriched molecular subtype will be randomized 1:1 to either:		
Treatment Arm A: Trastuzumab in combination with lapatinib: lapatinib 1000 mg PO once daily plus trastuzumab (loading dose of 8 mg/kg) followed by the maintenance dose of 6 mg/kg IV q3weekly. Weekly trastuzumab (loading dose of 4 mg/kg) followed by maintenance dose of 2 mg/kg IV weekly is acceptable at the discretion of the investigator. Patients with hormone receptor (ER and/or PgR)-positive, HER2-positive MBC randomized to this arm will be required to be treated with an	HER2-Enriched - Lapatinib (1000mg) +Trastuzumab (6mg per kg)+AI	1st column HER2-Enriched LAP+TRAS+AI Arm A

aromatase inhibitor of the investigator's choice.		
Treatment Arm B: Trastuzumab in combination with chemotherapy: trastuzumab (loading dose of 8 mg/kg) followed by maintenance dose of 6 mg/kg IV q3weekly plus chemotherapy of the investigator's choice. Weekly trastuzumab (loading dose of 4 mg/kg) followed by maintenance dose of 2 mg/kg IV weekly is acceptable at the discretion of the investigator. An aromatase inhibitor may be used for hormone receptor-positive, HER2-positive patients at the discretion of the investigator.	HER2-Enriched - Trastuzumab (6mg per kg)+Chemo +AI	2nd column HER2-Enriched TRAS+ CHEM+AI Arm B
COHORT 2 - Non-HER2-Enriched contains HER2+ patients with luminal A, luminal B and basal-like molecular subtypes		
Treatment Arm C: Trastuzumab in combination with lapatinib: lapatinib 1000 mg PO once daily plus trastuzumab (loading dose of 8 mg/kg) followed by the maintenance dose of 6 mg/kg IV q3weeks. Weekly trastuzumab (loading dose of 4 mg/kg) followed by maintenance dose of 2 mg/kg IV weekly is acceptable at the discretion of the investigator. Patients with hormone receptor (ER and/or PgR)-positive, HER2-positive MBC randomized to this arm will be required to be treated with an aromatase inhibitor of the investigator's choice.	Non-HER2-Enriched - Lapatinib (1000mg) +Trastuzumab (6mg per kg)+AI	3rd column Non-HER2-Enriched LAP+TRAS+AI Arm C

The patients' **actual treatment** will be derived from exposure data. If a patient's actual treatment is the same as the assigned treatment, then actual treatment is the assigned treatment. If a patient receives a study treatment that is different from the assigned treatment for the entire time of treatment, then actual treatment is the different treatment (the treatment actually received).

2.1.2.2 Study day

For randomized subjects, the date of randomization will be used as the reference date for efficacy and the date of the first dose will be used as the reference date for safety; for subjects in Arm C, the date of the first dose will be used as the reference date for both efficacy and safety.

For safety measures and other measurements (demographics, disease history, medical history, etc.), if the date of interest occurs on or after the safety reference date then the safety study day will be calculated as (date of interest - safety reference date) + 1. If the date of interest occurs before the safety reference date then the safety study day will be calculated as (date of interest – safety reference date). There is no safety study day 0.

For efficacy, If the date of interest occurs on or after the efficacy reference date then efficacy study day will be calculated as (date of interest - efficacy reference date) + 1. If the date of interest occurs prior to the efficacy reference date then efficacy study day will be calculated as (date of interest – efficacy reference date). There is no efficacy study day 0.

2.1.2.3 Duration

Durations (e.g., the duration of an adverse event, duration of exposure, etc.) are calculated as the stop date minus the start date plus one.

For elapsed time (e.g., the time since initial diagnosis):

- If the reference date is on or after the event date, then the elapsed time is the reference date minus the event date + 1.
- If the reference date is before the event date then the elapsed time is the reference date minus the event date.

When reporting time-to-event durations in months, divide the number of days by 30.4375; to report in weeks divide the number of days by 7; to report in years divide the number of days by 365.25. These algorithms for time to event return decimal numbers, and ignore the actual numbers of days in the months or years between start date and stop date. The "year" used in these algorithms is 365.25 days long, and the "month" is one twelfth of that year.

For converting all other durations (e.g., duration of adverse events, duration of exposure, age) to weeks, months or years use the following:

- To report the duration in weeks divides the number of days by 7.
- To report the duration in months use:
  (YEAR(stop date + 1) - YEAR(start date)) x 12 + (MONTH(stop date + 1) - month(start date) - 1) + (DAY(stop date + 1) > = DAY(start date))
- To report the duration in years use:
  intck('year', start date, stop date + 1) - (month(stop date + 1) < month(start date) or (month(stop date + 1) = month(start date) and day(stop date + 1) < day(start date)))

The algorithms above for age and duration return whole numbers for months and years, accurately accounting for the actual numbers of days in the months or years between the start date and the stop date.

2.1.2.4 Baseline and change from baseline

Baseline value will be defined as the last value before the first dose of study drug administered.

For the purposes of calculating change from baseline data, the screening or "Pre-dose/ Day 1" value will be used as baseline in all calculations. If this is missing then the change from baseline will also be missing. If more than one screening / pre-dose assessment is performed, then the one closest to and preceding start of treatment will be used for summaries.

For records occurring after baseline: change from baseline = (visit value) – baseline value. If either the baseline or visit value is missing, the change value is set to missing.

2.1.2.5 Follow-up period

In case of disease progression during the treatment period, the patient will be followed-up for 30 days for safety evaluation.

In case of study treatment discontinuation for any reasons other than disease progression, the patient will be followed-up for safety and efficacy assessments until disease progression, new anticancer therapy, death, withdrawal of consent or end of study, whichever comes first.

The safety data, up to 30 days after last dose, will be included in the summary statistics. For efficacy, any follow-up data will be included in the PFS analysis. Listings of safety and efficacy data will be provided to include all time-points captured.

2.1.2.6 Time windows and unscheduled visits

Time windows will not be applied. Visit based summaries will include scheduled assessments only and will be summarized according to the planned assessment time. The only exceptions to this are summaries of Left Ventricular Ejection Fraction (LVEF) change from baseline and, laboratory parameters where maximum toxicity grades and values outside laboratory reference ranges from any post-screening visit (include all scheduled and unscheduled visits) will be used.

Multiple assessments on ECG collected within a short timeframe (within 30 minutes) will qualify as a triplicate. The mean of one scheduled and the first two unscheduled measurements within 30 minutes will be considered as a point estimate of observed replicates. For ECG result (interpretation) the worst result should be taken (i.e., worst of the recorded of the following choices: abnormal – clinically significant, abnormal- not clinically significant, normal).

All assessments (scheduled and unscheduled) will be included in listings.

2.2 Analysis sets

The Intent-to-Treat (ITT) Population as defined in the protocol, also referred to as Full Analysis Set (FAS), will be used for the analysis of demography, baseline disease characteristic and secondary efficacy data and will consist of all patients who were: 1) randomized to study Arm A or Arm B, or (2) assigned to study Arm C, regardless of whether they actually received study treatment.

The Evaluable Population (EP) as defined in the protocol will be used for the primary analysis of change in biomarkers at disease progression, and will consist of all patients who actually

received the study treatment and have both baseline and progression tumor biopsies available, with evaluable data for at least 1 biomarker.

The Safety Population as defined in the protocol also referred to as Safety Set, will be used to assess clinical safety and tolerability and will consist of all patients who took at least 1 dose of study medication. This population will be based on the actual treatment received by the patients.

2.2.1 Subgroup of interest

As the study has a small sample size, there will be no subgroup analyses.

2.3 Patient disposition, demographics and other baseline characteristics

Unless otherwise stated, all tables and listings in this section will be based on the FAS, and all summaries and data listings will use treatment labels as specified in Section 2.1.4.1.

2.3.1 Patient disposition

Summary of patients screened by country and center was reported in the primary CSR. A summary of the number of patients in each of the analysis populations described was also produced for the primary CSR. These summary tables and corresponding listings will not be repeated. A summary of patient status and reason for study treatment discontinuation will be provided. This display will show the number and percentage of patients who discontinued from study treatment, and the primary reasons for treatment discontinuation which will be presented in the order they are displayed in the eCRF.

2.3.2 Protocol deviations

A summary and listing of all patients with important protocol deviations will be provided.

2.3.3 Demographic and baseline characteristics

Demographic and baseline characteristics, which were reported in the primary CSR, will not change and therefore will not be repeated for the final analysis. The summary and listing of the past and current medical conditions will be updated.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

Treatment compliance will be assessed for lapatinib, trastuzumab, chemotherapy, and aromatase inhibitor (AI).

Planned and actual treatments will be displayed for all study treatments under listing of exposure to that treatment.

A summary of overall compliance for lapatinib, oral chemotherapy and AI based on the exposure data will be produced. Percentage overall compliance will be summarized using the

mean, standard deviation, minimum, median, and maximum. In addition, percentage overall compliance will be categorized by <80%, 80%-120%, and >120%.

Overall compliance (%) = [total cumulative actual dose / (duration of study treatment x planned dose)] x 100, where duration of study treatment is last dose date - first dose date + 1.

Planned dose for lapatinib is 1000mg daily.

In addition, summaries of exposure to study medication (trastuzumab, lapatinib, chemotherapy and AI) and dose modifications (e.g., number of dose reductions, escalations, and interruptions) will further characterize compliance. These analyses are described in Section 'Extent of Exposure'.

2.4.2 Extent of exposure

The mean daily dose, cumulative dose, duration of treatment period, calculated as the number of days between start of treatment and end of treatment inclusive (i.e., treatment stop date minus treatment start date + 1), and relative dose intensity (%) will be summarized for lapatinib, chemotherapy and AIs.

The exposure to trastuzumab by visit will be summarized.

Exposure to study treatment will also be listed.

The following calculations will be performed for lapatinib, chemotherapy and AI exposure:

[1] Daily dose = Cumulative dose/ (Last administration date - First administration date + 1).

[2] Cumulative dose = Sum of dosing intervals [(Last date in interval - First date in interval +1) x Dose level in interval].

[3] Duration of exposure (week) = (Last administration date - First administration date + 1)/7.

[4] Relative dose intensity (%) = 100 x Dose intensity /Planned dose intensity.

Following definition will be used in the computation:

- Planned Treatment Duration (days): it is the planned time between two consecutive administration (planned cycle duration)
- Dose Intensity (mg/week) (DI): Cumulative dose (mg) / Treatment duration (weeks)
- Planned DI (mg/week): (PDI) Cumulative planned dose (mg) / Planned Treatment duration (weeks)
- Relative DI (%) (RDI): 100 x DI (mg/w) / PDI (mg/w). An RDI of 100% indicates that the drug was administered at the right dose within the planned timeframe (e.g. every 28 days).

Dose reductions for lapatinib will be summarized by number of reductions, reasons for reductions, and number of reductions at each visit. Dose interruptions for lapatinib will be summarized by number of interruptions, reasons for the interruption, duration of interruption (days), and number of interruptions at each visit.

Dose delays for trastuzumab will be summarized by number of delays, reasons for the delay, delay duration (days), and number of delays at each visit. These data will also be listed.

A summary of study treatment status will be provided. This display will show the number and percentage of patients who are ongoing or discontinued study treatment and a summary of the primary reasons for discontinuation of study treatment. Reasons for study treatment discontinuation will be presented in the order they are displayed in the eCRF. A listing of study treatment discontinuation will be generated; this will include last dose date, and reasons for study treatment discontinuation.

2.4.3 Prior, concomitant and post therapies

A listing of concomitant medications will be presented. The concomitant medications taken by each patient and the relationship between the Anatomical-Therapeutic-Chemical (ATC) Level 1, preferred term and verbatim text will be displayed. All medication verbatim text will be coded using the most recent version of WHO Drug Dictionary to obtain appropriate ATC Level 1 classification and ingredients. A summary and listing of prior and concomitant medications by ATC Level 1 and ingredients by treatment arm will be provided.

2.5 Analysis of the primary objective

The analysis of the primary objective will be based on the evaluable population (or EP or Evaluable Set). All supportive analyses of primary endpoint will be based on the FAS Set.

2.5.1 Primary endpoint

The primary endpoint of this study is the change in expression profile of genes and/or proteins within each arm on a set of biomarkers (more than 700 immune genes on nanostring panel) associated with immunomodulation between the pre-treatment biopsy and disease progression biopsy.

2.5.2 Statistical hypothesis, model, and method of analysis

2.5.2.1 Biomarker data normalization

mRNA gene expression panel derived from Nanostring platform will be transformed and normalized by the vendor before it can be used in the analysis. The method is described as follows: There are 5 housekeeping genes recommended by the lab: ACTB, MRPL19, PSMC4, RPLP0 and SF3A1. The normalization factor will be computed as the geometric mean of housekeeping genes across samples. The expression of each gene will be divided by this normalization factor.

2.5.2.2 Data summaries

Tissue biopsies are taken at two time points: baseline and disease progression, whereas whole blood is taken every 9 weeks. Tumor infiltrating lymphocytes (TILs), [redacted], ER & PR will be tested on the tissue biopsies at both time points (at baseline and at disease progression). Flow cytometry data will be generated on whole blood samples that were taken every 9 weeks.

For all tissue biopsy samples, data will be summarized at the baseline and on-progression visits.

- The expression of genes and/or proteins will be for both baseline and on-progression visits.

- The fold-change of expression of genes and/or proteins will be only on-progression

Flow cytometry data will not be summarized due to insufficient number of subjects. The flow cytometry data will be listed.

Changes in TILs in paired pre-treatment tumor biopsies and those obtained at disease progression for each arm will be summarized in tables that include sample size, geometric mean, SD, geometric %CV, median, minimum and maximum. In addition, known markers CD4 and CD8A for immune cell populations and immune microenvironment will be analyzed comparing their levels at disease progression to the baseline to assess how the treatments affect the immune response at the level of the tumor. In case there are more than one data points at baseline, mean value will be used as baseline. TILs changes from pre-treatment to disease progression will also be listed.

The fold-change from baseline to disease progression of gene expression values will be summarized for each arm. The summary tables will present sample size, geometric mean, SD, geometric %CV, median, 95% CI of the median, minimum and maximum. The fold-changes from pre-treatment to disease progression for the gene expression will also be listed.

Since the panel of genes is very large, results will be displayed only for a selected subset of these genes.

Ranking and selection of genes

The gene subset will be selected based on the largest changes (increase or decrease) from baseline to disease progression, within each treatment arm separately. Biomarker data should have sample collection date from which it should be possible to derive the study day/visit.

The following algorithm will be implemented within each arm separately, to select up to 100 genes for display:

1. The significant fold-change will be flagged.
   a. if the 95% CI of the median fold change of a gene does not include 1, Flag.significant=1, otherwise Flag.significant =0.
2. Flag the genes with increase from baseline
   a. If median fold change >1, flag.gt=yes, otherwise flag.gt=no.
3. Flag the genes with decrease from baseline
   a. If median fold change <1, flag.lt=yes, otherwise flag.lt=no.
4. Rank those genes with increase from baseline (i.e. flag.gt=yes)
   a. rank first by the significance flag (Flag.significant)
   b. then rank in order of median fold change value, from largest to smallest fold-change values.
   c. Select the genes list based on those that have ranks from 1 to 50. If there are less and 50 genes that increase from baseline, select all.
5. Rank those genes with decrease from baseline (i.e. flag.lt=yes)
   a. rank first by the significance flag (Flag.significant)
   b. then rank in order of median fold change value, from smallest to largest fold-change value.

c. Select the genes list based on those that have ranks from 1 to 50. If there are less and 50 genes that decrease from baseline, select all.

6. The complete summary table (for each arm) will be presented for the selected genes.

2.5.3 Handling of missing values/censoring/discontinuations

Missing data will not be imputed for biomarkers.

2.5.4 Supportive analyses

No additional analysis will be conducted due to a small number of subject sample collected.

2.6 Analysis of the secondary objective

All secondary analyses are intended to describe the summary statistics by the treatment arm. All the changes will be evaluated within each individual arm. No pooling of patients over arms will be done for secondary analyses.

2.6.1 Secondary endpoint

2.6.1.1 Progression free survival (PFS)

PFS is defined as the interval of time (in months) between the date of randomization and the earlier of date of disease progression and the date of death due to any cause. Disease progression will be based on the assessments by the investigator using RECIST 1.1, and is based on FAS.

The date of documented disease progression will be defined as the date of radiological disease progression as assessed by the investigator based on imaging data only. The date of documented disease progression will be taken from the Response Assessment page. This should be the date of the lesion evaluation corresponding to the documented disease progression. If there are multiple scan dates for a given assessment then the first date in the series of scans that shows progression of the disease should be used. It should be noted that in the cases of patients who have PD determined by non-radiological evidence (e.g. lab data or symptomatic progression); it will not be considered an event. Such patients' data should be checked for any later radiological evidence if this available and only censored in line with the following rules.

- The date of death should be taken from the Record of Death page. Death on study due to any cause will be included.
- If a patient has neither progressed by radiological assessment nor died, then PFS will be censored at the date of the last radiological scan as assessed by the investigator. This applies also in the case of patients who are withdrawn from the study prior to progression or death. Such patients will be censored at the last radiological scan assessment.
- Patients who receive subsequent anti-cancer therapy (including radiotherapy and surgeries with anti-cancer intent) in the absence of documented disease progression will be censored at the initiation of the subsequent therapy if scans or progression do not occur within 14 days of the initiation of therapy.

For patients who receive subsequent anti-cancer therapy, the following rules will apply:

- If the start date of the anti-cancer therapy is partial (i.e. either missing the day but has the month and year available or missing both day and month), the imputation rules described in the SAP for the primary CSR will be applied.
- If such therapy is started without evidence of documented disease progression, then PFS will be censored at the date of the last radiological assessment that is no later than the date of initiation of subsequent anti-cancer therapy if the scan assessment does not document PD. If there are multiple scan dates for a given assessment then PFS will be censored at the first date of the series of scan dates.
- If the start of the subsequent anti-cancer therapy plus 14 days is after the date of documented disease progression (defined as the first date if there are multiple dates in a series of scans) then the patient will be considered to have had a progression event and not censored.
- If the start of the subsequent anti-cancer therapy plus 14 days is prior to the date of documented disease progression, then the patient will be censored on the date of the previous scan.
- If initiation of subsequent anti-cancer therapy plus the 14-day window is prior to documented disease progression and the anti-cancer therapy start date is between the first and last date of a set of multiple scans, and the date of subsequent disease progression is also between the first and last date of a set of multiple scans, then PFS will be censored at the first scan date of the previous assessment.

If a patient has only a baseline visit or does not have a date of radiological scan that is no later than the date of initiation of anti-cancer therapy, PFS will be censored at the date of randomization.

2.6.1.2 Handling of missing values/censoring/discontinuations for PFS

Missing values will not be imputed. The censoring of PFS is described in Section 2.6.1.1.

2.6.1.3 PFS estimates

PFS will be summarized using Kaplan-Meier curves. Percentage of patients with progression free status at monthly intervals will be obtained from the Kaplan-Meier estimate of the progression free survival function. This will include at a minimum the following time points: 3 months, 6 months, 9 months and 12 months. Time point estimates will be provided as is reasonable based on the available data. Approximate 95% confidence limits will be calculated, based on Greenwood's formula for the standard error of the Kaplan-Meier estimate. For each treatment arm, the Kaplan-Meier estimates for the median PFS time, the first and third quartiles will be presented along with approximate 95% CI's.

PFS time will be fitted to a Cox proportional hazard model with treatment as a covariate. The hazard ratios between the treatment and the corresponding 95% CI will be presented.

2.6.1.4 Overall response rate (ORR)

The response for each visit is 'according to the investigator assessment of response'. Use the response for each visit from CRF, not re-compute the response at each time-point using target, non-target lesions measurement. However, the best overall response is computed. Table 2-2

presents the overall response at an individual time-point for all possible combinations of tumor responses in target and non-target lesions with or without the appearance of new lesions for patients with measurable disease at baseline according to RECIST.

Table 2-2 Evaluation of Overall Response for Patients with Measurable Disease at Baseline

Target Lesions	Non-Target Lesions	New Lesions	Overall Response
CR	CR or NA	No	CR
CR	Non-CR/Non-PD or NE	No	PR
PR	Non-PD or NA or NE	No	PR
SD	Non-PD or NA or NE	No	SD
NE	Non-PD or NA or NE	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

CR=complete response, PR = partial response, SD=stable disease, PD=progressive disease, NA= Not applicable, and NE=Not Evaluable

Table 2-3 presents the overall response at an individual time-point for all possible combinations of tumor responses in non-target lesions with or without the appearance of new lesions for patients with non-measurable only disease at baseline.

Table 2-3 Evaluation of Overall Response for Patients with Non-Measurable Disease at Baseline

Non-Target Lesions	New Lesions	Overall Response
CR	No	CR
Non CR/Non PD	No	Non CR/Non PD
NE	No	NE
Unequivocal PD	Yes or No	PD
Any	Yes	PD

CR=complete response, PD=progressive disease, and NE=Not Evaluable

Note:

- Patients with a global deterioration of health status requiring discontinuation of treatment without objective evidence of disease progression at that time should be classified as having "symptomatic deterioration." Objective response status is determined by evaluations of disease burden. Every effort should be made to document the objective progression even after discontinuation of treatment.
- In some circumstances, it may be difficult to distinguish residual disease from normal tissue. When the evaluation of CR depends on this determination, it is recommended that the residual lesion be investigated (fine needle aspirate/biopsy) to confirm the CR.

The best overall response is the best response recorded from the start of the treatment until disease progression/recurrence and will be determined programmatically based on the investigators assessment of response at each time-point.

- To be assigned a status of SD, follow-up disease assessment must have met the SD criteria at least once after randomization at a minimum interval of 56 days (8 weeks).
- If the minimum time for SD is not met, best response will depend on the subsequent assessments. For example if an assessment of PD follows the assessment of SD and SD does not meet the minimum time requirement the best response will be PD. Alternatively, patients lost to follow-up after an SD assessment not meeting the minimum time criteria will be considered not evaluable.

Confirmed ORR is defined as the percentage of patients achieving either a confirmed CR or PR. Tabular summary will be produced for all patients (patients with measurable and non-measurable disease).

- Confirmed CR - at least two determinations of CR at least 4 weeks apart before disease progression;
- Confirmed PR - at least two determinations of PR or better at least 4 weeks apart before progression (and not qualifying for a CR).

Confirmed ORR will be calculated from the investigator's assessment of response. Patients with unknown or missing response will be treated as non-responders; i.e. they will be included in the denominator when calculating the percentages. Further if a patient only has a baseline measurement, per RECIST version 1.1, they should be classified as NE and appear in the table as such, so that the sum of the individual response categories (CR, PR, SD, PD, NE) sum up to the big N, where N is the number of patients in the denominator (i.e. the number of patients in the FAS in each treatment arm).

ORR will be summarized along with the exact 95% confidence interval (CI) for each treatment arm.

The data on target, non-target and new lesion will be listed.

2.6.1.5 Clinical benefit rate (CBR)

CBR is defined as the percentage of patients achieving either a confirmed CR or PR tumor response at any time or maintaining SD for at least 24 weeks while on study, according to the investigator assessment of response per RECIST 1.1 criteria. The SD should last at least 24 weeks. All patients who achieve a CR, PR or SD will be included in these analyses, including any patients with non-measurable disease who achieve a CR, (according to the qualifying criteria in Table 2-3), which has a duration of at least 26 weeks.

An estimate of the proportion of patients who experience clinical benefit will be summarized and analyzed similarly as described above for ORR. Patients with unknown or missing response will be treated as non-responders (i.e. not CR, PR or SD<24 weeks). Patients who have withdrawn from the study prior to completing 6 months of response assessments will still be counted in this analysis in the denominator, and hence considered as patients who do not experience clinical benefit in the numerator.

2.6.1.6 Handling of missing values/censoring/discontinuations

Missing values on the response are considered as non-responses for ORR. Please refer to censoring rules for PFS in Section 2.6.1.1.

2.6.1.7 Derivation of lesion summary variables

When any target lesion data was updated at any visit and/or if scheduled or unscheduled assessments were entered out of order, the derived fields should also be updated using the 'Recalculate' button on target lesion assessment CRF page. In order to ensure the accuracy of the derived fields, the values of the following variables will be derived at each visit for the analysis dataset.

- Sum of diameters
- Baseline sum of diameters
- Percent change in the sum of diameters of target lesions, taking as reference the baseline [% change from baseline = (sum of lesion - sum of lesion at baseline) / sum of lesion at baseline x 100]
- Smallest sum of diameters (nadir)
- Absolute change in the sum of diameters of target lesions, taking as reference the smallest sum of diameters (nadir) recorded at previous assessments including baseline [absolute change from nadir = sum of lesion - nadir (smallest prior)]
- Percent change in the sum of diameters of target lesions, taking as reference the smallest sum of diameters (nadir) recorded at previous assessments including baseline [% change from nadir = (sum of lesion - nadir sum of lesion) / nadir sum of lesion x 100]Association between biomarkers and PFS

Analysis of association between biomarkers and PFS will not be performed due to insufficient sample size from early closure of the study.

[redacted]

2.8 Safety analyses

The Safety population as defined in Section 2.2 will be used for the analysis of safety data.

Safety data, up to 30 days after last dose, will be included in the summary. All safety data will be listed.

2.8.1 Adverse events (AEs)

All AE summaries will be based on AEs collected during the on-therapy period (from start of first dose of study medication to last dose of study medication) plus 30 days after last dose. All AEs and SAEs will be listed for all patients with pre- and post-treatment flag (indicated by study day).

An overview summary of AEs collected during the study, including counts and percentages of patients with any AE, AEs related to study treatment, AEs leading to permanent discontinuation

of study treatment, SAEs, SAEs related to study treatment, fatal SAEs, and fatal SAEs related to study treatment will be produced.

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables for on-treatment AEs that are not serious with an incidence greater than 5% and on-treatment SAEs and SAEs suspected to be related to study treatment will be provided by system organ class and preferred term on the safety set population.

If for the same patient, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- A single occurrence will be counted if there is ≤ 1 day gap between the end date of the preceding AE and the start date of the consecutive AE
- More than one occurrences will be counted if there is > 1 day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE/SAE suspected to be related to study treatment /non SAE has to be checked in a block e.g., among AEs in a ≤ 1 day gap block. If at least one SAE is occurring, then one occurrence is counted for that SAE.

Adverse events (AEs) will be graded according to the CTCAE, Version 4.0. Adverse events will be grouped by system organ class (SOC) and preferred term (PT) using the Medical Dictionary for Regulatory Affairs (MedDRA). MedDRA version 20.1, which was used for coding AEs for the primary CSR, will continue to be used for the final CSR. Note that incorrect MedDRA version 21.0 was cited in the SAP for primary CSR.The summary of on-treatment AEs, treatment-related AEs, AEs leading to permanent discontinuation of study treatment and AEs related to treatment will be presented for each treatment arm.

A study treatment-related AE is defined as an AE for which the investigator classifies the relationship to study treatment as “Yes”. The summary table will include events with the relationship to study treatment as ‘Yes’ or missing.

Adverse events will also be presented by treatment arm and the maximum NCI CTCAE toxicity grade reported.

For all AE tables, AEs will be sorted by SOC. SOC will be presented in descending order from the SOC with the highest total incidence to the SOC with the lowest total incidence. A SOC will not be presented when the overall total incidence is zero. Within each SOC preferred terms will be presented in descending order from the preferred term with the highest total incidence to the preferred term with the lowest total incidence. Where two or more preferred terms within an SOC have the same total incidence, they will be sorted alphabetically. Terms relating to either diarrhea or rash will be collapsed and presented as one preferred term in the overall summaries of adverse events. If the SOC for rash terms is ‘Infections’, it will be recoded to ‘Skin and Subcutaneous Tissue’ SOC to ensure correct reporting. A further summary of the individual terms will also be provided.

A summary of the most common adverse events will be presented by treatment arm. A preferred term will be considered amongst the most common adverse events if its occurrence is $\geq 10\%$ (before rounding). This summary will differ from the one detailed above in that only preferred terms will be presented, not SOCs.

Separate listings will be provided to support each AE summary. Listings identifying patient numbers for the individual adverse events will be provided. Relationship of AE system organ class, preferred terms and verbatim text will also be listed.

2.8.2 Adverse events of special interest / grouping of AEs

The following AEs have been defined as AEs of special interest (AESI): rash-related events, diarrhea related events, hepatobiliary events, cardiac related events and interstitial lung disease.

The AESI groups and the preferred terms are shown in Table 2-4.

Table 2-4 Adverse Events of Special Interest

AE of Special Interest	AE Preferred Term
Rash Related Events	Acne Dermatitis Dermatitis acneiform Eczema Erythema Exfoliative rash Photosensitivity reaction Rash Rash generalized Rash macular Rash maculo-papular Rash papular Rash pruritic Skin ulcer Rash pustular*
Diarrhea Related Events	Diarrhea Frequent bowel movements
Cardiac Related Events	Acute left ventricular failure Acute right ventricular failure Cardiac failure Cardiac failure acute Cardiac failure chronic Cardiac failure congestive Chronic left ventricular failure Chronic right ventricular failure Cor pulmonale Cor pulmonale acute Cor pulmonale chronic Ejection fraction abnormal Ejection fraction decreased Left ventricular dysfunction Left ventricular failure Right ventricular failure Ventricular dysfunction

	Ventricular failure
Hepatobiliary events	Acute hepatic failure Alanine aminotransferase Alanine aminotransferase abnormal Alanine aminotransferase increased Ammonia abnormal Ammonia increased Aspartate aminotransferase Aspartate aminotransferase abnormal Aspartate aminotransferase increased Autoimmune hepatitis Bilirubin conjugated abnormal Bilirubin conjugated increased Bilirubin urine Blood alkaline phosphatase Blood alkaline phosphatase abnormal Blood alkaline phosphatase increased Blood bilirubin Blood bilirubin abnormal Blood bilirubin increased Blood bilirubin unconjugated Blood bilirubin unconjugated increased Cholestatic liver injury Cytolytic hepatitis Gamma-glutamyltransferase Gamma-glutamyltransferase abnormal Gamma-glutamyltransferase increased Hepatic encephalopathy Hepatic enzyme abnormal Hepatic enzyme increased Hepatic failure Hepatic function abnormal Hepatic infiltration eosinophilic Hepatic necrosis Hepatic steatosis Hepatitis Hepatitis acute Hepatitis cholestatic Hepatitis fulminant Hepatitis toxic Hepatobiliary disease Hepatocellular injury Hepatotoxicity Hyperammonaemia Hyperbilirubinaemia Hypertransaminasaemia

	Jaundice Jaundice cholestatic Jaundice hepatocellular Liver disorder Liver function test Liver function test abnormal Liver injury Subacute hepatic failure Transaminases Transaminases abnormal Transaminases increased
Interstitial lung disease	Acute interstitial pneumonitis Lung Infiltration Interstitial lung disease Pneumonitis

* SOC will be recoded to 'Skin and subcutaneous tissue disorders'

The number and percentage of patients having at least one event within each AESI group will be summarized by treatment group and toxicity grade (all grades, grades ≥ 3). The AESIs will also be summarized by SOC and PT within each treatment group. Patient listing of AESIs will also be produced.

2.8.3 Deaths and serious adverse events (SAEs)

SAEs will be summarized for each treatment arm by system organ class and preferred term.

The summary of on-treatment SAEs, treatment-related SAEs, 'SAEs leading to permanent discontinuation of study treatment', 'On-treatment deaths and Serious Adverse Events' and 'Fatal AEs related to study treatment' will be presented for each treatment arm.

All deaths recorded on the Record of Death eCRF page will be summarized and listed. The primary reason for death will be coded.

Further, a summary detailing all on-treatment deaths will be produced. This will consist of all deaths that occurred prior to the discontinuation of study treatment, with the allowance of a window of 30 days. Hence, on-treatment deaths should be selected as any deaths occurring from day 1 of first dose received until last dose received + 30 safety follow-up days.

The number of deaths resulting from SAEs suspected to be related to study treatment and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

Separate listings will be provided to support each non-fatal SAE and fatal SAE summary.

2.8.4 Adverse events leading to discontinuation of study treatment and/or withdrawal from the study and other significant adverse events

All adverse events leading to permanent discontinuation of the study treatment will be summarized and listed. Summary of adverse events leading to dose modification, dose delay, and dose reduction, dose escalation, early stopping of infusion and infusion interruption but with infusion completion will also be provided.

2.8.5 Clinical laboratory evaluations

Laboratory parameters will be summarized using descriptive statistics. Shifts from baseline will be categorized as Normal range High, within Normal range and Normal range Low. Additionally, lab toxicities will be displayed by Maximum toxicity grade. Laboratory toxicity grades will be based on worst toxicity grade.

Hematology and clinical chemistry data will be summarized at each scheduled assessment, and at study treatment discontinuation. Summaries of worst-case grade increase from baseline grade will be provided for all the lab tests that are gradable by CTCAE v4.0. These summaries will display the number and percentage of patients with a maximum post-baseline grade increasing from their baseline grade. Any increase in grade from baseline will be summarized along with any increase to a maximum grade of 3 and any increase to a maximum grade of 4. Missing baseline grade will be assumed as grade 0. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia.

For lab tests that are not gradable by CTCAE v4.0, summaries of worst-case changes from baseline with respect to normal range will be generated. Decreases to low, changes to normal or no changes from baseline, and increases to high will be summarized for the worst case post-baseline. If a patient has a decrease to low and an increase to high during the same time interval, then the patient is counted in both the "Decrease to Low" categories and the "Increase to High" categories.

Supporting listings for haematology and chemistry laboratory tests will be produced.

A supporting listing of laboratory data for patients with abnormalities of CTC grade 3 or 4 will be provided. Unless otherwise specified, the denominator in percentage calculation at each scheduled visit will be based on the number of patients with non-missing value at each particular visit.

Laboratory test results will be reported in International System of Units (SI).

Summary and listing of hematology and serum chemistry parameters will be presented in SI unit. The listing of local labs by country, center, sex, and age will also be displayed with local lab name, address, original unit of lab parameters and normal ranges.

2.8.6 Analyses of liver function test

Hepatobiliary lab abnormalities will be summarized and listed for any patients with lab abnormalities for aspartate aminotransferase (AST), alanine aminotransferase (ALT), total bilirubin and alkaline phosphatase (ALP) according to the categories listed in Table 2-5 :

Table 2-5 Hepatobiliary Lab Abnormality Criteria

Category
Possible Hy's Law: >3x ULN ALT & ≥ 2x ULN BIL [1] & (<2x ULN ALP or ALP missing)
>3x ULN ALT & ≥ 2x ULN BIL [1]
Hepato-cellular injury [2]

> 3x ULN ALT or AST
> 5x ULN ALT or AST
> 8x ULN ALT or AST
> 20x ULN ALT or AST
> 3x ULN ALT
> 5x ULN ALT
> 8x ULN ALT
> 20x ULN ALT
> 3x ULN ALT and ≤ 3x ULN ALT Baseline
≥ 2x ULN BIL [1] and < 2x ULN BIL Baseline
≥ 3x ULN ALP and < 3x ULN ALP Baseline
≥ 2x ULN BIL [1]
≥ 3x ULN ALP

ALT: alanine aminotransferase; AST: aspartate aminotransferase; ALP: alkaline phosphatase; BIL: Total Bilirubin; ULN=Upper Limit of Normal.

[1] If Direct Bilirubin is available, then Direct Bilirubin must also be >35% when Total Bilirubin is ≥2x ULN in order to satisfy the criteria. Bilirubin value can occur up to 28 days on or after ALT value.

[2] Hepato-cellular injury is defined as R=[Patient's ALT / ULN of ALT] / [Patient's ALP / ULN of ALP] ≥5

Possible Hy's law cases are defined as any elevated ALT>3×ULN, total bilirubin≥2×ULN and ALP<2×ULN/missing. Total bilirubin≥2×ULN can be within 30 days following the ALT elevation and if direct bilirubin is available on the same day, it must be ≥ 35% of total bilirubin. ALP<2xULN/missing means the criteria is satisfied unless the ALP is >2xULN at any time of bilirubin elevation within the 30 day window.

2.8.7 Liver toxicity

All liver events meeting either stopping or follow-up criteria as defined in Section 5.9.1 of the protocol will be listed. Results of liver chemistry, imaging, and other parameters as required will be listed. Liver pharmacokinetic measurements obtained for patients experiencing a respective hepatobiliary event will be listed with time elapsed since onset of the event.

The assessment of causality of a hepatobiliary event will be done with the use of the data collected for Roussel Uclaf Causality Assessment Method (RUCAM). RUCAM scores will not be calculated. The data collected will be presented in a listing.

2.8.8 Other safety measures

2.8.8.1 LVEF

Absolute change from baseline in LVEF will be summarized at each scheduled assessment time and in the worst case post-baseline. Only the post-baseline assessments that used the same method (ECHO or MUGA) as the baseline assessments will be used to derive the change from

baseline. The change from baseline (LVEF decrease from baseline value) will be categorized as below.

- No change or any increase
- Any decrease
- >0-<10% decrease
- 10-19% decrease
- ≥20% decrease
- ≥10% decrease and ≥ LLN
- ≥10% decrease and below LLN
- ≥20% decrease and ≥ LLN
- ≥20% decrease and below LLN

LVEF results will also be listed with patient level details including absolute change from baseline.

2.8.8.2 ECG

The ECG parameters include: QT, HR, and QTcB. QTcB considers both: QTcB = machine read and QTcBc = calculated manually from QTc using Bazett formula.

Summaries of 12-lead ECG data will include data from scheduled assessments only. In case of a scheduled record occurred with multiple unscheduled records, the scheduled record with the first two unscheduled assessments occurring within 30 minutes form a triplicate. The mean of the triplicate will be used as a point estimate instead for the triplicate. All data will be reported according to the nominal visit date for which they were recorded (i.e., no visit windows will be applied). Missing value will not be imputed. If unexpected value occurs (such as text where numeric is expected) then it will be set to missing.

The standard 12-lead ECG results at baseline and post-baseline visits will be summarized as normal, abnormal - clinically significant, or abnormal - not clinically significant. All data on ECG will be listed. For ECG result (interpretation) the worst result should be taken (i.e., worst of the recorded of the following choices: abnormal – clinically significant, abnormal- not clinically significant, normal).

Clinical notable (CN) criteria will be assessed through the descriptive displays listed below. In all cases, post baseline refers to any post-baseline assessment including unscheduled visits. If unscheduled visits are associated with a triplicate, then the mean of the triplicate will be used.

Notable criteria for QT and QTcB:

- Increase >30 to <=60 ms
- Increase >60 ms
- New >450 to <=480 ms
- New >480 to <=500 ms
- New >500 ms

Notable criteria for HR:

- Increase >25% and HR >100 bpm
- Decrease >25% and HR <50 bpm

'Increase/decrease' refers to change from baseline; 'new' is also relative to baseline. For example, if a subject has QTcB of 450 ms at baseline and 451 ms at a post-baseline visit, then this post-baseline measure from the subject meets this criteria: 'New >450 to <=480 ms'.

Listing of cardiovascular events (CV) occurred during the study period will be displayed.

2.8.8.3 Vital signs

Vital sign assessments include weight, blood pressures (diastolic and systolic), heart rate, and temperature. Change from baseline in vital signs will be summarized by treatment arm. Descriptive statistics will be shown for baseline, the result at each IP dose, and at the EOS visit, and the change from baseline to each IP dose result and at the EOS visit.

Post-baseline vital signs will be defined as clinically notable if they meet the criterion value listed in the table below. The number and percentage of subjects who meet the criterion value and the number and percentage of subjects who meet both the criterion value and the change from baseline criterion will be summarized for any time post-baseline for each vital sign parameter. Data from unscheduled visits will be included.

Table 2-6 Criteria for Treatment Emergent Clinically Notable Vital Signs

Vital Sign Parameter	Flag	Criterion Value	Change from Baseline
SBP (mmHg)	High (CH)	≥ 180	Increase of ≥ 20 mmHg
	Low (CL)	≤ 90	Decrease of ≥ 20 mmHg
DBP (mmHg)	High (CH)	≥ 105	Increase of ≥ 15 mmHg
	Low (CL)	≤ 50	Decrease of ≥ 15 mmHg
Heart Rate (bpm)	High (CH)	≥ 120	Increase of ≥ 15 bpm
	Low (CL)	≤ 50	Decrease of ≥ 15 bpm

The listing of vital sign data will be displayed.

2.9 Pharmacokinetic endpoints

Liver pharmacokinetic data from any patient who experiences a liver adverse event as defined by the protocol will be listed.

2.10 Patient-reported outcomes

Up to approval of protocol amendment 2, the multidimensional assessment of fatigue (MAF) and the FACT-B questionnaire were not collected anymore. The previously collected data were listed in the primary CSR.

2.11 Biomarkers

The biomarker analyses are the primary analyses. These are defined in Section 2.5. If any further biomarker analyses are required, they will be defined according to the results of the primary

analysis. These analyses will be the presented in a separate analysis plan and study report if such analysis is conducted.

2.12 Interim analysis

Other than the analysis for the primary CSR, there is no interim analysis.

3 Sample size calculation

Sample size is not based on statistical power to detect difference between the treatment arms. The number of patients in the study is based on the patient accrual rate. Eighty-seven patients were screened, 42 patients have been randomized (32 patients with HER2- enriched and 10 patients with HER2- Non-enriched population). A decision was made to close the study early. Recruitment was closed on 31 March 2017.

4 Change to protocol specified analyses

No change from protocol specified analysis was made.

5 Reference

Edge SB, Byrd DR, Compton CC, et al., eds.: AJCC Cancer Staging Manual. 7th Ed. New York, NY: Springer, 2010, pp 347-76.

Gaby Danan and Rolf Teschke. RUCAM in Drug and Herb Induced Liver Injury: The Update. International Journal of Molecular Sciences; 2016; 17: 14.