SIGNATURE INFORMATION

Document: 1200-0034--tsap-revision-1

Document No.: T12-1008-02

Title LUX-Lung 6; A randomised, open-label, phase III study of BIBW 2992 versus chemotherapy as first-line treatment for patients with stage IIIB or IV adenocarcinoma of the lung harbouring an EGFR activating mutation

SIGNATURES (ELECTRONICALLY OBTAINED)

Meaning of Signature:	Signed by:	Date signed: (GMT)
Author-Trial Clinical Pharmacokineticist		12/5/2012 13:11:25
Author-Trial Clinical Monitor		12/5/2012 13:20:13
Approval-Project Statistician		12/5/2012 13:22:25
Author-Trial Medical Writer		12/5/2012 13:52:00
Author-Trial Statistician		12/5/2012 14:45:47

SIGNATURE INFORMATION (continued)

Document 1200-0034--tsap-revision-1

Document No.: T12-1008-02

Title LUX-Lung 6; A randomised, open-label, phase III study of BIBW 2992 versus chemotherapy as first-line treatment for patients with stage IIIB or IV adenocarcinoma of the lung harbouring an EGFR activating mutation

SIGNATURES (ELECTRONICALLY OBTAINED)
(There are no entries on this page if there are up to seven signatures.)

Meaning of Signature:	Signed by:	Date signed: (GMT)

Trial Statistical Analysis Plan

T12-1008-02

BI Trial No.:	1200.34
Title:	LUX-Lung 6; A randomised, open-label, phase III study of BIBW 2992 versus chemotherapy as first-line treatment for patients with stage IIIB or IV adenocarcinoma of the lung harbouring an EGFR activating mutation Including protocol amendments 1 [U10-3034-01-AM1], 2 [U10-3034-01-AM2] and 3 [U10-3034-01-AM3].
Investigational Product:	BIBW 2992
Responsible trial statistician:	Tel: Fax:
Date of statistical analysis plan:	05 December 2012 Revised
Version:	Revised

1. TABLE OF CONTENTS

LIST OF TABLES

2. LIST OF ABBREVIATIONS

Term	Definition / description
AE	Adverse event
ANCOVA	Analysis of Covariance
AUC	Area under the curve
BRPM	Blinded Report Planning Meeting
CI	Confidence interval
CR	Complete Response
CTC	Common Terminology Criteria
CTP	Clinical Trial Protocol
CTR	Clinical Trial Report
ECOG	Eastern Cooperative Oncology Group
EGFR	Epidermal Growth Factor Receptor
EMA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
FAS	Full analysis set
GFR	Glomerular filtration rate
HRQOL	Health-Related Quality of Life
ICH	International Conference on Harmonisation
IPV	Important Protocol Violation
LOCF	Last observation carried forward
LVEF	Left Ventricular Ejection Fraction
MedDRA	Medical Dictionary for Regulatory Activities
MQRM	Medical Quality Review Meeting
NE	Not Evaluable
NN	Non-CR/Non-PD
O*C	Oracle Clinical
OS	Overall Survival
PFS	Progression Free Survival
PK	Pharmacokinetics
PPS	Per protocol set
PR	Partial Response

Term	Definition / description
PT	Preferred term
Q1	Lower quartile
Q3	Upper quartile
QLQ	Quality of Life Questionnaire
RECIST	Response Evaluation Criteria in Solid Tumours
SA	Statistical analysis
StD	Standard deviation
SD	Stable Disease
SOC	System Organ Class
TOC	Table of contents
TSAP	Trial statistical analysis plan

3. INTRODUCTION

As per ICH E9, the purpose of this document is to provide a more technical and detailed elaboration of the principal features of the analysis described in the protocol, and to include detailed procedures for executing the statistical analysis of the primary and secondary variables and other data.

This Trial Statistical Analysis Plan (TSAP) assumes familiarity with the Clinical Trial Protocol (CTP), including Protocol Amendments. In particular, the TSAP is based on the planned analysis specification as written in CTP Section 7 “Statistical Methods and Determination of Sample Size”. Therefore, TSAP readers may consult the CTP for more background information on the study, e.g., on study objectives, study design and population, treatments, definition of measurements and variables, planning of sample size, randomization.”

SAS® Version 9.2 will be used for all analyses.

4. CHANGES IN THE PLANNED ANALYSIS OF THE STUDY

The following changes to those statistical methods described in the CTP and its subsequent amendments were made:

- The CTP describes that the effect of Afatinib on PFS compared with standard chemotherapy will be tested at the one-sided, 0.025 significance level. This is identical to the effect of Afatinib being tested at the more commonly used two-sided, 0.05 (5%) significance level if the treatment effect is in favour of Afatinib. In order to aid in the interpretation of this study two-sided p-values will therefore be presented throughout.
- Further details to those given in the CTP have been provided for the timing of the statistical analyses; refer to Section 7 for further details.
- An additional row (3d) was added to the Table 7.4: 1 to cover a scenario that had not been considered at the time of writing the CTP.
- The CTP specifies that patients who are randomised but never receive randomised treatment will be censored on the day of randomisation, for both PFS and overall survival (OS) analyses. In order to handle all randomised patients the same, irrespective of whether they received randomised treatment or not, the above censoring rule will not be considered. Refer to Table 7.4: 1 for a complete list of censoring rules.
- The CTP specifies that if a Health-Related Quality of Life (HRQOL) assessments is missed, but followed by another assessment which shows deterioration then the time to deterioration will be defined as the mid-point of the two available assessments immediately before and after the missed assessments. As it was thought that this planned imputation was likely to be a rare occurrence as well as being unnecessarily complicated it will not be performed. Refer to Section 7.5.2.4 for further details.
- The time to deterioration in body weight and Eastern Cooperative Oncology Group (ECOG) performance status endpoints have been removed as descriptive summaries of changes from baseline will suffice. Refer to Section 7.5.2 for further details.
- The CTP states that adverse events will be compared over a period of time equivalent to six courses of chemotherapy, such a comparison was included due the expected difference in length of treatment exposure between the two treatment groups. This analysis has been replaced by exposure adjusted adverse event incidence rate summaries which have the benefit of not excluding any treatment emergent adverse events from summary. Refer to Section 7.8.1 for further details.
- The list of laboratory parameters to be focussed upon detailed in the CTP was revised; refer to Section 7.8.2 for further details.

5. ENDPOINTS

5.1 PRIMARY ENDPOINT

The primary efficacy endpoint will be progression free survival (PFS) calculated as the time from randomisation to the occurrence of disease progression or death, whichever occurs first. Disease progression will be evaluated according to Response Evaluation Criteria in Solid Tumours (RECIST) 1.1 criteria.

5.2 SECONDARY ENDPOINTS

5.2.1 Key secondary endpoints

- Objective response to study treatment, defined as complete response (CR) or partial response (PR) as determined by RECIST 1.1. Time to objective response, defined as the time from randomisation to the first recorded objective response. Duration of objective response will be measured from the time of first objective response to the time of progression or death, whichever occurs first (or date of censoring for PFS).
- Disease control, defined as a patient with an objective response or stable disease (SD). Duration of disease control will be measured from randomisation to the time of progression or death, whichever occurs first (or date of censoring for PFS).
- Overall survival (OS) calculated as the time from randomisation to death.

5.2.2 Other secondary endpoints

- Tumour shrinkage measured as the minimum sum of target lesion diameters recorded after randomisation.
- Change from baseline in body weight.
- Change from baseline in ECOG performance status.
- HRQOL as measured by the European Organisation for Research and Treatment of Cancer (EORTC) core Quality of Life Questionnaires (QLQ-C30) and lung cancer specific questionnaire module (QLQ-LC13).

6. GENERAL ANALYSIS DEFINITIONS

6.1 TREATMENTS

For reporting purposes, all randomised patients will be classified into either 'Afatinib 40' or 'Ge1000+Cis75' as randomised. For efficacy analyses patients will be analysed as randomised. For safety analyses, treated patients will be analysed according to the initial treatment taken.

The following study periods based on actual start and stop dates of study treatment administration are defined:

- Pre-screening: for any data recorded prior to day of first signing informed consent.
- Screening: day of first signing informed consent to day prior to starting study treatment.
- On-treatment: day of first administration of study treatment to day of last administration of study treatment.
- Post-treatment: day after last administration of study treatment to the 28th day after last administration of study treatment.
- Post-study: 29th day after last administration of study treatment.

For safety summaries data recorded up to 28 days after last administration of study treatment will be considered as on-treatment (i.e. the actual on-treatment and post-treatment periods defined above will be combined into one 'on treatment' analysing treatment).

6.2 IMPORTANT PROTOCOL VIOLATIONS

No per protocol set (PPS) analysis will be performed for this study; however patients with potentially important protocol violations (IPVs) will be documented. The following list of potentially IPVs will be used; note that this is a working list and may not be finalised until the final Blinded Report Planning Meeting (BRPM) prior to database lock.

Table 6.2: 1 Important protocol violations

Category/ Code		**Description of Violation**	**Comment/Example**
A		**Entrance criteria not met**	
	A1	Diagnosis including disease staging and measurability not as per protocol.	Refers to IN1 and IN3, also check oncological history and baseline tumour measurement details from investigator.
	A2.1	Prior treatment for NSCLC does not meet entrance criteria	Refers to EX1, EX2, EX3, and EX24 also check previous therapy details.

Table 6.2: 1 Important protocol violations - continued

Category/ Code		Description of Violation	Comment/Example
	A2.2	Genotypic tumour characteristics do not meet entrance criteria	Refer to IN2 also check tumour biopsy details.
	A2.3	Baseline imaging taken outside of the protocol defined window	Date of baseline imaging of target and non-target lesions must be no more than 28 days from starting study treatment; and not after starting study treatment.
	A3	Laboratory values do not meet entrance criteria	Refers to EX11, EX12, EX13, EX14 and EX15, also check against screening laboratory values.
	A4	Chronic diarrhoea or other pre-existing event that might exacerbate expected Afatinib events	Refers to EX7, EX10 and EX23 also check baseline conditions.
	A5	Other deviation from entrance criteria	Refers to IN4, 5 and 6 and EX4, 5, 6, 8, 9, 16, 17, 18, 19, 20 and 22. Also check baseline demographics, oncological history, ECG, LVEF, pregnancy test, ECOG score and baseline conditions.
B		**Informed consent**	
	B1	Patient's written informed consent not available	Refers to IN7. Date of written informed consent 1 and 2 must be available.
	B2	Patient's written informed consent too late	Date of written informed consent must be on or before screening visit 1. Date of written informed consent 2 must and on or before screening visit 2.
	B3	Procedure performed prior to written informed consent	Study procedure performed prior to patient giving informed consent.
C		**Trial medication and randomisation**	

Table 6.2: 1 Important protocol violations - continued

Category/ Code		Description of Violation	Comment/Example
	C1	Incorrect trial medication taken	For Afatinib patients: starting dose of Afatinib treatment not 40mg. Afatinib dose not paused, reduced or increased according to Section 4.1.4 of the CTP. For chemotherapy patients: chemotherapy dose calculated incorrectly, >6 courses given or <6 courses without good reason.
	C2	Randomisation not followed	Patient not treated according to randomisation. Patient randomised into incorrect stratum.
	C3	Non-compliance	Check medication compliance details for extreme non-compliance only.
D		**Concomitant medication**	
	D2	Prohibited medication use	Review concomitant medications for prohibited medication use, refer to Section 4.2.2 of the CTP.
E		**Missing primary endpoint data**	
	E4	Non-evaluable by RECIST criteria during the entire trial	Note individual non-evaluable timepoints will be described as part of a description of the accuracy of the PFS determination.
G		**Trial specific**	
	G1	Non-adherence to safety related withdrawal criteria	Check for patients that are continuing when they should have been withdrawn from treatment/trial according to the criteria detailed in Section 6.3 of the CTP.

6.3 PATIENT SETS ANALYSED

- Randomised set:
  This patient set includes all patients randomised to receive treatment, whether treated or not. This set of patients will be used for the primary evaluation of efficacy.

- Treated set:
  This patient set includes all patients who were dispensed with and documented to have taken at least one dose of study medication. This set of patients will be used for the evaluation of safety.

6.5 POOLING OF CENTRES

This section is not applicable because centre is not included in any of the statistical models; it is not included to avoid over-fitting the model.

6.6 HANDLING OF MISSING DATA AND OUTLIERS

In general missing efficacy data will not be imputed and all reasonable efforts will be taken during the study to obtain such data. Patients with unknown vital status, missing tumour imaging data and missing HRQOL assessments will be censored for time to event analyses; further details are provided in Section 7.

Missing or incomplete adverse event (AE) dates are imputed according to BI standards ['Handling of missing and incomplete AE dates' 001-MCG-156_RD-01 (1)].

Missing data and outliers of PK data are handled according to BI standards ['Non-compartmental Pharmacokinetic/Pharmacodynamic Analyses of Clinical Studies' 001-MCS-36-472 RD-01 (2)].

6.7 BASELINE, TIME WINDOWS AND CALCULATED VISITS

Baseline values will be the measurements taken most recently prior to first administration of the study drug.

Study day will be calculated relative to the date of the first administration of study drug. The day prior to first administration of study drug will be 'Day -1' and the day of first administration of study drug will be 'Day 1'; therefore 'Day 0' will not exist.

All data collected at a visit level will be presented according to the actual protocol visit it was measured at. Visits will be labelled as Screening Visit 1 (SV1), Screening Visit 2 (SV2), Course 1 Visit 1 (C1 V1), C1 V2, C2 V1, C2 V2, C3 V1, C4 V1 … CX V1, End of Treatment (EOT), Follow-up 1 (FUP1) … FUP X and Observation Period 1 (OP1) … OP X. Data obtained at unplanned/additional visits will be listed and suitably labelled.

For the presentation of tumour response data which will follow a calculated visit approach based on the protocol specified tumour imaging schedule. Imaging will be performed at 6-weekly intervals up until Week 48 and in 12-weekly intervals thereafter from the date of randomisation; images will be slotted to Week 6, 12, 18, … XX based on their relative day (from randomisation) and using a ± 3, ±6 week window as appropriate (images taken in the

first 3 weeks from randomisation will be slotted to Week 6). If two or more images for a patient are assigned to one interval then the last assessment will be used to ensure progressive disease is not missed.

7. PLANNED ANALYSIS

For end-of-text tables, the set of summary statistics is: N / Mean / StD / Min / Median / Max. For appendix tables, the set of summary statistics is: N / Mean / StD / Min / Q1 (lower quartile) / Median / Q3 (upper quartile) / Max.

Tabulations of frequencies for categorical data will include all possible categories and will display the number of observations in a category as well as the percentage (%) relative to the respective treatment group. Percentages will be rounded to one decimal place. The category missing will be displayed only if there are actually missing values. Percentages will be based on all patients in the respective patient set whether they have non-missing values or not.

The primary analysis will take place when 217 patients have progressed or died, as assessed by independent review. The time point of the data cut-off (extract of data from BI database) for sending the latest data packages (radiological scans and completed dossiers) for independent review will be determined by the progression and death data from the investigator assessments. Based upon the discrepancy observed (on blinded data) between the number of events from the independent review (ongoing during the trial) and the investigator assessments the time point for observing 217 independent review events will be estimated. All data up to and including this time point (including response data from patients who did not progress) will be sent for independent review. The estimation of this time point is likely to lead to some deviation, though relatively small, from the planned 217 independent review events. The trial will only be un-blinded after the independent review data has been received back, loaded into the database and considered final.

A second analysis, primarily for assessing OS, will be performed when approximately 209 have died. The timing of this analysis may be adjusted to coincide with future regulatory submissions or if an update is required at the time a decision is being made regarding regulatory approval. All data collected after this will be reported in a revision to the CTR.

7.1 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Descriptive statistics using standard summary tables for the randomised set of patients are planned for this section of the report.

7.2 CONCOMITANT DISEASES AND MEDICATION

Descriptive statistics using standard summary tables for the randomised set of patients are planned for this section of the report.

7.3 TREATMENT COMPLIANCE

Descriptive statistics using standard summary tables for the treated set of patients are planned for this section of the report. A summary of whether patients took their treatment according to the protocol and whether they missed any doses (Afatinib only) will be produced for each planned visit. In addition a summary of overall percentage compliance (Afatinib only) will be produced using visit dates and the total number of doses missed during the study.

7.4 PRIMARY ENDPOINT

The primary endpoint of this study is PFS, defined as the time (months) from the date of randomisation to the date of disease progression, or to the date of death if a patient died earlier. The date of progression for the primary analyses will be determined by an independent central imaging review incorporating both radiologist and oncologist reviews, blinded to treatment assignments. The primary endpoint analysis will be performed on the randomised set of patients.

A stratified log-rank test (two-sided, 0.05 significance level) will be used to test the effect of Afatinib on PFS compared with standard chemotherapy. The test will be stratified by EGFR mutation group (L858R, Del 19 or Other), according to the patient data entered in the CRF.

A Cox proportional hazards model, stratified by EGFR mutation group will be used to estimate the hazard ratio and 95% confidence interval (CI) between the two treatment groups. Kaplan-Meier estimates and 95% CIs (using Greenwood's standard error estimate) will be tabulated at 3-monthly time points and will include a comparison of the treatment groups using a z-test (approximation of the normal distribution). Kaplan-Meier curves for the two treatment groups will also be produced.

The assumption of proportional hazards within each EGFR mutation group and the homogeneity of the hazard ratio across the EGFR mutation groups will be checked descriptively.

The rules to determine whether or not patients have had a PFS event (progression or death) along with the date of event or date of censoring (for those with no event) are specified in Table 7.4: 1. These rules will be applied to the independent central review data prior to analysis. A summary table of the reasons for censoring will be produced by treatment group. Event and censoring rates will also be presented at each planned imaging time point.

Table 7.4: 1 Rules to determine events and censoring for PFS

Rule #	Situation	Outcome (event or censored)	Date of PFS event or censoring
1	No baseline tumour assessment (no death before second scheduled assessment)	censored	Date of randomisation
2	Progressed from central imaging (no missed radiologic assessments)	event	Date of PD
3a	Non-PD from central imaging[1], death before next scheduled assessment	event	Date of death
3b	Non-PD from central imaging[1], one missed assessment, death or progression after date of missed assessment, but before a second scheduled assessment	event	Date of PD or death
3c	Non-PD from central imaging[1], more than one consecutive missed assessment, death or progression after date of second missed assessment	censored	Date of last imaging before missed assessment
3d	Non-PD from central imaging[1], more than one consecutive missed assessment, non-PD according to imaging after missed assessments	censored	Date of last non-PD imaging
4	New anti-cancer medication before progression or death	censored	Date of last imaging before new anti-cancer medication
5	Death before the scheduled date of first imaging	event	Date of death
6a	No imaging performed post-baseline, patient dies between first and second scheduled assessments	event	Date of death
6b	No imaging performed post-baseline, patient dies after second scheduled assessment	censored	Date of randomisation
6c	No imaging performed post-baseline, vital status is unknown or patient known to be alive	censored	Date of randomisation
7	Alive and not progressed from central imaging (no missed assessments)	censored	Date of last imaging

[1] - From the last assessment at which CR, PR or SD was assessed.

7.5 SECONDARY ENDPOINTS

Secondary endpoints are split into 'key' and 'other' endpoints. If the difference between the treatment groups is statistically significant then formal statistical testing will be performed on the key secondary endpoints. In order to protect the overall type I error rate a closed testing

procedure will be employed, whereby each key secondary endpoint will only be formally analysed if the previous endpoint was found to be statistically significant. The key secondary endpoints will be analysed according to the order they are presented below.

All analyses of secondary endpoints will be produced on the randomised set of patients.

7.5.1 Key secondary endpoints

7.5.1.1 Best RECIST tumour assessment

Each patient will be given a best overall response of CR, PR, SD, Non-CR/Non-PD (NN), PD or Not Evaluable (NE) based on the independent central imaging data, following the RECIST (1.1). This will be based on all responses taken from the start of treatment until the start of any new anti-cancer therapy.

Objective response

An objective response is defined as a best overall response of CR or PR. A logistic regression model, stratified by EGFR mutation group will be used to compare the objective response rate between the two treatment groups. This analysis will be considered as the first key secondary endpoint analysis. Objective response rates will be presented with exact 95% Clopper-Pearson CIs.

For those patients with an objective response, time to response measured from the date of randomisation to the date of first response will be summarised by the planned imaging time points. A descriptive summary of the duration of response (months), measured from the first date of response to progression or death will also be produced. Kaplan-Meier curves for the two treatment groups will also be produced for the duration of response, where applicable patients will be censored as for the PFS primary analysis.

The RECIST guidelines state that confirmation of response is not required for Phase III studies that have PFS as the primary endpoint; hence the above summaries will be produced without the requirement for confirmation. However, for completeness all the above summaries and analyses will be repeated with the requirement for confirmation, these will be considered secondary in nature.

Disease control

Disease control is defined as a best overall response of CR, PR, SD or NN. A logistic regression model, stratified by EGFR mutation group will be used to compare the disease control rate between the two treatment groups. This analysis will be considered as the second key secondary endpoint analysis. Disease control rates will be presented with exact 95% Clopper-Pearson CIs.

A descriptive summary of the duration of disease control (months), measured from the date of randomisation to progression or death will also be produced. Kaplan-Meier curves for the two treatment groups will also be produced for the duration of disease control, where applicable patients will be censored as for the PFS primary analysis.

As for the objective response summaries, analyses of disease control will be repeated with the requirement for confirmation, these will be considered secondary in nature.

7.5.1.2 Overall survival

OS (months), defined as the time from the date of randomisation to the date of death will be formally analysed twice and will be considered as the third and final key secondary endpoint. The first analysis will be performed at the time of the primary PFS analysis and the second at a time when more information is available on OS. To preserve the overall 0.025, one-sided alpha level, a Haybittle-Peto stopping boundary will be used (p-value<0.0001) for the first analysis.

Patients for whom there is no evidence of death at the time of analysis will be censored on the date that they were last known to have been alive.

A stratified log-rank test will be used to test the effect of Afatinib on OS compared with standard chemotherapy. The test will be stratified by EGFR mutation group. A Cox proportional hazards model, stratified by EGFR mutation group will be used to estimate the hazard ratio and 95% confidence interval (CI) between the two treatment groups. Kaplan-Meier estimates and 95% CIs (using Greenwood's standard error estimate) will be tabulated at 3-monthly time points and will include a comparison of the treatment groups using a z-test (approximation of the normal distribution). Kaplan-Meier curves for the two treatment groups will also be produced.

7.5.2 Other secondary endpoints

7.5.2.1 Tumour shrinkage

Tumour shrinkage for each patient, measured (via central imaging review) as the minimum sum of target lesion diameters after randomisation, will be compared for the two treatment groups. An analysis of covariance (ANCOVA) for the minimum sum of diameters, with the baseline sum of diameters fitted as a covariate and EGFR mutation group fitted as a factor will be performed.

In addition waterfall plots of the maximum percentage decrease from baseline will be presented for each treatment group.

7.5.2.2 Body weight

The lowest and last body weight recorded on treatment or during follow-up will be summarised along with the change in weight from baseline.

7.5.2.3 ECOG performance status

A shift table of the worst (highest) and last ECOG performance status category recorded on treatment or during follow-up by baseline category will be produced.

7.5.2.4 Health-related quality of life

HRQOL questionnaires as measured by the EORTC QLQ-C30 and QLQ-LC13 at baseline, during treatment and follow-up will be included in the analysis. All scoring of the symptom scales/items will follow the EORTC scoring algorithm.

The analyses will focus on cough, dyspnoea and pain; specifically:

- Cough: Q1 - How much did you cough? (QLQ-LC13)
- Dyspnoea: Composite of: Q3 - Were you short of breath when you rested? Q4 - Were you short of breath when you walked? Q5 - Were you short of breath when you climbed stairs? (QLQ-LC13). Individual item from QLQ-C30, Q8 - Were you short of breath?
- Pain: Composite of: Q9 - Have you had pain? Q19 - Did pain interfere with your daily activities? (QLQ-C30). Individual items from QLQ-LC13, Q10 - Have you had pain in your chest? Q11 - Have you had pain in your arm or shoulder? Q12 - Have you had pain in other parts of your body?

For each of the summary scales and items measuring the above the treatment groups will be compared for the following:

Distribution of patients improved, stable or worsened

Improvement will be defined as a score that improves from baseline by at least 10 points (on the 0-100 point scale) at anytime during the study. If a patient has not improved, worsening will be defined as a 10 point worsening at anytime during the study. Otherwise, a patient will be considered as stable.

The number and percentage of patients falling into each of these three categories will be summarised by treatment group. A logistic regression model, stratified by EGFR mutation group will be used to compare the distribution of patients improving/not improving across the two treatment groups.

Time to deterioration

Time to deterioration (months) is defined as the time from randomisation to an increase (worsening) from the baseline score of at least 10 points on the 0-100 point scale. Patients that die during the period of interest with no evidence of a deterioration will be considered to have deteriorated at the time of death.

This time to event data will be analysed and summarised using the same methodology as for the primary efficacy endpoint, refer to Section 7.4. The results of the analyses will be displayed using Forest plots.

Patients with no deterioration (including those with disease progression) will be censored at the last available HRQOL assessment date; patients with no post-baseline assessments will be censored on the day of randomisation.

Change in scores over time

Changes in scores over time will be assessed using longitudinal models; these will be mixed-effects growth curve models with the average profile over time for each endpoint described by a piecewise linear model adjusted for the fixed effect of EGFR mutation group. The models will allow the slope to change at 3, 6, 12 and 18 weeks. Each model will also include the two random effects of intercept and slope (time from randomisation). The area under the estimated growth curve (AUC) up to the median follow-up time will be calculated as a summary measure for each treatment group; this will be divided by the median follow-up time so that it can be interpreted as the mean HRQOL score up to the median follow-up time. The treatment effect will be estimated as the average difference between the treatment group means scores, together with a 95% confidence interval and associated p-value based on a t-statistic with degrees of freedom calculated using the Kenward-Roger method. The results of the analyses will be displayed using Forest plots.

In addition, all single items and subscales (functional and symptom) from both questionnaires will be analysed in a similar fashion to summarise the impact of therapy over the entire profile of the measures, and to examine the consistency of component items with the composite measures.

Note as higher scores represent a 'better' level of functioning; deterioration in scales/items related to functioning will be defined as a decrease from baseline score of at least 10 points.

Finally, the usage of cough, dyspnoea and pain medication along with the incidence of cough and dyspnoea AEs will be described.

7.7 EXTENT OF EXPOSURE

Total treatment time (days and number of courses) will be calculated for each patient; off-drug periods due to non-compliance or toxicity prior to permanent discontinuation will be included as treatment time. For Afatinib, 21 days will be considered a treatment course. In addition the total treatment time will be summed over all patients and transformed to patient years. Standard descriptive summaries, by treatment group, of these data will be provided for the treated set of patients.

A Kaplan-Meier plot showing the number of patients at risk (exposed) during the study treatment period will also be produced.

Further summaries will also be produced for the Afatinib treatment group only:

- Treatment time (days) broken down by each dose level (50 mg, 40 mg, 30 mg and 20 mg).
- Number and proportion of patients on each dose level over time.
- Time to first dose reduction and duration (days) of off-drug periods prior to first dose reduction.
- Kaplan-Meier plot of time from first dose of study medication to the first dose reduction

7.8 SAFETY ANALYSIS

All safety analyses will be performed on the treated set.

7.8.1 Adverse events

Unless otherwise specified, the analyses of AEs will be descriptive in nature. All analyses of AEs will be based on the number of patients with AEs and not on the number of AEs.

Furthermore, for analysis of AE attributes such as duration, severity and CTC grade multiple AE occurrence data on the CRF, will be collapsed into AE episodes provided that all of the following applies:

- The same MedDRA lowest level term was reported for the occurrences.
- The occurrences were time-overlapping or time-adjacent (time-adjacency of 2 occurrences is given if the second occurrence started on the same day or on the day after the end of the first occurrence).
- Treatment did not change between the onset of the occurrences or treatment changed between the onset of the occurrences, but no deterioration was observed for the later occurrence

For further details on summarization of AE data, please refer to the guideline 'Handling and summarization of adverse event data for clinical trial reports and integrated summaries' [001-MCG-156 (7)].

The main analysis of adverse events will be based on the concept of treatment emergent adverse events. That means that all adverse events occurring between first drug intake until 28 days after last drug intake will be assigned to the randomised treatment. All AEs occurring outside of this period will be handled according to the treatment definition rules detailed in Section 6.1 and will only be listed unless stated otherwise.

An overall summary of adverse events will be presented. This summary will exclude the rows ‘Severe AEs’, ‘Significant AEs’ and ‘Other significant AEs’ but will include additional rows for ‘AEs leading to dose reduction’ and ‘AEs by highest Common Terminology Criteria (CTC)’ grade.

The frequency of patients with adverse events will be summarised by highest CTC grade (grades 3, 4, 5 and all grades to allow both treatment groups to fit on one page), treatment, primary system organ class (SOC) and preferred term (PT) for each of the following AE tables:

- All AEs.
- Drug related AEs.
- AEs leading to dose reduction.
- AEs leading to treatment discontinuation.
- Drug related AEs leading to treatment discontinuation
- Serious AEs.
- Drug related serious AEs.
- AEs leading to death.
- AEs in Afatinib patients who dose escalate from 40 mg to 50 mg.

All tables will be sorted by SOC according to the standard sort order specified by the European Medicines Agency (EMA); PTs will be sorted by frequency (within SOC).

The above tables except for AEs leading to death will be repeated with the project defined grouping of AE terms (rash, stomatitis, ocular effects, lip effects, nail effects and fatigue). Details of the project defined groupings are defined in the technical TSAP. In these tables the grouped AEs will replace the original PTs for all AEs that are included within the grouped term. The grouped AE categories will then be tabulated along with all remaining MedDRA PTs, sorted by descending frequency. A reference table presenting all of the project defined groupings and MedDRA PTs within each grouping will also be produced.

Additional AE tables will be produced for AEs of special interest (diarrhoea, and the project defined groupings of rash, renal insufficiency, leukopenia and neuropathy), providing further details on highest CTC grade, action taken with study drug and time to first onset of AE.

A supportive analysis of adverse events will be performed using treatment exposure adjusted incidence rate tables. These will be produced for the project defined grouping of AE terms (rash, stomatitis, ocular effects, lip effects, nail effects and fatigue) in a similar manner to that described above. The following tables will be produced for all AEs and for all AEs with a CTC grade ≥3:

- All AEs.
- Drug related AEs.
- Serious AEs.

Specifications for the calculation of incidence rates, rate ratios and CI for rate ratios are detailed in the technical TSAP (also refer to 8).

7.8.2 Laboratory data

The analyses of laboratory data will be descriptive in nature and will be based on BI standards ['Display and Analysis of Laboratory Data' 001-MCG-157 (9)]. CTC grades will be applied to laboratory parameters using the current BI oncology standard as detailed in the document 'Conversion of laboratory parameters to CTCAE grades within BI' (10).

Descriptive statistics of all laboratory values by visit will be provided including changes from baseline. Frequency tables of transitions relative to the reference range and of possible clinically significant abnormalities will be produced. For those parameters that have CTC grading possible clinically significant abnormalities are defined as those laboratory values with a CTC grade ≥ 2 that have had an increase of ≥ 1 grade from baseline. For those parameters for which no CTC grade has been defined standard BI project definitions will be used to decide on clinical significance. Further frequency tables will show the transition of CTC grade from baseline to worst value and last value on treatment.

Summaries will be produced of laboratory data recorded prior to treatment, on-treatment and 28 days after last study drug intake.

The focus of the laboratory data analysis will be on the following laboratory parameters:

- Low values: haemoglobin, white blood cells, neutrophils, lymphocytes, platelets, potassium, sodium and glomerular filtration rate.

- High values: creatinine, asparate aminotransferase, alanine aminotransferase, bilirubin, alkaline phosphatase and creatinine phosphokinase.

7.8.3 Vital signs

Only descriptive statistics are planned for this section of the report.

A summary table of left ventricular ejection fraction (LVEF) will be produced for the final and minimum on-treatment values along with the corresponding percentage change from baseline. The number of patients with a significant LVEF event will also be presented; a significant event is defined as a decrease of ≥20% from baseline that is also below the lower limit of normal for the particular site (50% will be used if the lower limit of normal is missing).

7.8.4 ECG

A descriptive summary of QTcF values will be produced by each scheduled visit along with the change from baseline. All other ECG data will be listed.

7.8.5 Others

The relationship between trough Afatinib concentration, categorised using the quartile values, will be explored for the following endpoints:

- Week 6 tumour shrinkage; measured as the absolute and percentage change in the sum of target lesion diameters from randomisation.
- PFS and OS.

Trough Afatinib concentrations will be taken as the values measured on Day 42 prior to study treatment intake; if this value is missing then the value taken on Day 29, or Day 21 if that is missing will be used.

8. REFERENCES

1. 001-MCG-156_RD-01, ‘Handling of missing and incomplete AE dates’, current version; IDEA for CON.

2. 001-MCS-36-472 RD-01, 'Non-compartmental Pharmacokinetic/Pharmacodynamic Analyses of Clinical Studies', current version; IDEA for CON)

3. R10-4999, P Korhonen and J Palmgren 'Effect modification in a randomized trial under non-ignorable non-compliance: an application to the alpha-tocopherol beta-carotene study', Applied Statistics (2002) 51, Part 1: 115-133

4. R10-5000, P Korhonen, N Laird and J Palmgren 'Correcting for non-compliance in randomized trials: An application to the ATBC study', Statistics in Medicine (1999). 18, 2879-2897

5. R10-4487, J Robins and D Finkelstein 'Correcting for Noncompliance and Dependent Censoring in an AIDS Clinical Trial with Inverse Probability of Censoring Weighted (IPCW) Log-Rank Tests', Biometrics (Sept 2000), 56, 779-788

6. R10-4484, J Robins 'Information recovery and bias adjustment in proportional hazards regression analyses of randomized trials using surrogate markers' Proceedings of the Biopharmaceutical Section, American Statistical Association (1993), 24-33

7. 001-MCG-156, ‘Handling and summarisation of adverse event data for clinical trial reports and integrated summaries’, current version; IDEA for CON.

8. R07-1386, Rothman K and Greenland S. Modern Epidemiology, 2nd ed. 1998

9. 001-MCG-157, ‘Display and Analysis of Laboratory Data’, current version, IDEA for CON.

10. BI Guidance document ‘Conversion of laboratory values to CTCAE grades within Boehringer Ingelheim’

10. HISTORY TABLE

Table 10: 1 History table

Version	Date (DD-Mmm-YY)	Author	Sections changed	Brief description of change
Revised	**05-Dec-12**		4	Removal of details regarding reduced assay set.
Revised	**05-Dec-12**		6.2	Revision of IPV category A2.3. Removal of IPV category E3.
Revised	**05-Dec-12**		6.3	Reduced assay set removed as not required for this study
Revised	**05-Dec-12**			
Revised	**05-Dec-12**		7.4	
Revised	**05-Dec-12**		8	Change to location and identification of reference 2.