BELIEF

ETOP 2-11

An open-label phase II trial of erlotinib and bevacizumab in patients with advanced non-small cell lung cancer and activating EGFR mutations

Bevacizumab and **ErL**otinib **I**n **EGF**R mut + NSCLC

A clinical trial of ETOP

Coordinated by Grupo Español de Cancer de Pulmón GECP

(Spanish Lung Cancer Group)

STATISTICAL ANALYSIS PLAN FOR FINAL ANALYSIS

SAP VERSION	1.2
SAP VERSION DATE	06 November 2015
TRIAL STATISTICIANS	ETOP: Urania Dafni, Marie Kassapian
	SLCG: Maria Sanchez Ronco
TRIAL PRIMARY INVESTIGATORS	Rafael Rosell
	Rolf Stahel
SAP AUTHOR	Xanthi Pedeli, Zoi Tsourti
APPROVED BY	Urania Dafni

TABLE OF CONTENTS

ABBREVIATIONS

AE	Adverse Event
AEG-1	Astrocyte Elevated Gene 1
AESI	Adverse Event of Special Interest
ALT	Alanine Transaminase
AP	Alkanine Phosphatase
AST	Aspartate Transaminase
BRCA1	Breast Cancer 1, early-onset
CR	Complete Response
CT	Computed Tomography
CTCAE	Common Terminology Criteria for Adverse Events
DC	Disease Control
DCR	Disease Control Rate
DR	Duration of Response
EGFR	Epidermal Growth Factor Receptor
FFPE	Formalin Fixed, Paraffin Embedded
INR	International Normalized Ratio
NSCLC	Non-Small Cell Lung Cancer
OR	Objective Response
ORR	Objective Response Rate
OS	Overall Survival
PFS	Progression-Free Survival
PR	Partial Response
RECIST	Response Evaluation Criteria In Solid Tumours
SAE	Serious Adverse Event
TTD	Time to treatment Discontinuation
TTF	Time to Treatment Failure
ULN	Upper Limit of Normal lab value

1. INTRODUCTION

1.1 Preface

Advanced non-small-cell lung cancer (NSCLC) patients harbouring epidermal growth factor receptor (EGFR) mutations (del 119 or L858R) show an impressive progression-free survival (PFS) between 9 and 14 months when treated with erlotinib. However, the presence of EGFR mutations can only imperfectly predict outcome. The leading hypothesis of the current project is that PFS could be influenced both by the pre-treatment EGFR T790M mutation and by components of DNA repair pathways. To investigate this hypothesis, a model of treatment is proposed whereby a combined treatment with erlotinib plus bevacizumab is supposed to benefit patients with EGFR mutations (single or with T790M). The trial encompasses two phase II substudies run in parallel, as shown in Figure 1. The first substudy follows Simon's two stage design and a decision can be reached at the first stage, while the second substudy follows Fleming's single stage design.
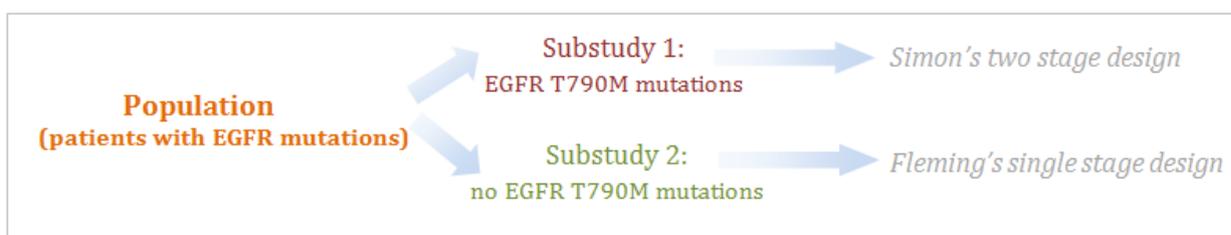
Figure 1. Trial's design flowchart

1.2 Hypotheses

Patients with EGFR mutations (single or with T790M) can attain a benefit with longer overall PFS when treated with erlotinib plus bevacizumab. More specifically, the hypotheses of interest are that the trial treatment with erlotinib plus bevacizumab will result in:

- *an increase of the median PFS to 18 months for patients with EGFR T790M mutation;*
- *a median PFS of approximately 18 months or more in patients without EGFR T790M mutations.*

1.3   Purpose of the present Statistical Analysis Plan

The purpose of the present Statistical Analysis Plan (SAP) is to provide an analytic and solid framework for the primary endpoint evaluation of the BELIEF study (final study), after all enrolled patients have reached one year of follow-up.

A comprehensive description of all the analyses planned for the study is provided in section "Planned Analyses".

More details on the analysis, in terms of methods, definitions and reporting, are described in detail in Sections 3-8.

The overall objectives and methods of the study, as dictated by its protocol, may be found in the Appendix.

2.   PLANNED ANALYSES

The total study duration is estimated to be approximately 48 months: 12 months' recruitment period including a 3 months run-in period, and at least 36 months treatment and follow-up period, since, according to the protocol, patients will be followed until death.

The analyses planned are:

- Periodic interim analyses for reviewing safety.

- Formal interim efficacy analysis for evaluating the designed stopping rule of the substudy 1 of the trial.

- Final analysis, including efficacy as well as safety analysis of the results of the clinical trial.

Interim analyses will be conducted, approximately every 6 months, **for reviewing the safety** of the study. ETOP's Independent Data Monitoring Committee (IDMC), consisting of scientific, clinical and biostatistical experts, is responsible for evaluating interim results. Only the IDMC is authorised to review interim efficacy and safety analysis results by subgroup of the study. The committee is responsible for reviewing accumulating safety data at regular intervals (adverse events), monitoring the overall study conduct (e.g. accrual rate), deaths and response, and making recommendations as to whether continuing study treatment and enrolment is scientifically and ethically appropriate (study is conducted in a manner that does not expose participants to undue risk).

Particularly for **substudy1 of EGFR with T790M mutation**, a **formal interim efficacy analysis** is planned, according to the trial design, to be conducted after the first 8 patients reach 12 months of follow-up (if not died or lost to follow-up). According to Simon's two-stage design, in the first-stage, if from the total of 8 patients, 4 or more patients reach 12 months without a progression-defining event, then the substudy proceeds so as to evaluate the 12-month PFS for a total of 35 patients (initially planned sample size for substudy1).

In the Simon's 2nd stage evaluation**,** the analysis of the (first) 35 evaluable patients reaching one-year follow-up will be performed: according to the design if at least 19 out of the 35 patients reach one-year from enrollment without a progression defining event, treatment

with the combination of erlotinib and bevacizumab would be considered promising in the subgroup of patients with EGFR T790M mutation.

3. GENERAL CONSIDERATIONS

3.1 Timing of Analysis

The final primary evaluation for all patients of the clinical trial (EGFR with or without T790M mutation), will be performed when at least one year from enrolment of all patients is reached.

3.2 Analysis Populations

Efficacy Cohort

The efficacy cohort will be consisted of all eligible patients enrolled in the trial (ITT population). The efficacy cohort will be used for the assessment of patient baseline characteristics and all efficacy endpoints. Efficacy analysis will be performed by substudy, as well as overall for all eligible, for efficacy, patients.

Safety Cohort

The safety cohort will encompass all patients who have received at least one dose of trial treatment.

3.3 Missing Data

Missing values will not be replaced by any statistics calculated over non-missing data.

3.4 Multiple Comparisons / Multiplicity

Apart from the adjustment made for the formal planned interim analysis for the subgroup of EGFR with T790 mutation patients, no additional adjustments will be made in this study.

3.5 Overall Cohort and Substudy Cohorts

All analyses will be performed overall for all patients in the efficacy or safety cohort, as well as for the 2 subgroups of patients corresponding to the two substudies, i.e. patients with EGFR with or without T790M mutation.

Secondary analysis based on subgroups of EGFR mutations Exon19/Exon21 will also be performed.

4. SUMMARY OF STUDY DATA

4.1 Primary and secondary endpoints

The primary endpoint according to the study protocol, is the **progression-free survival (PFS)** for all enrolled patients. PFS is defined as the time from the date of enrolment until documented progression or death, whichever occurs first.

Secondary endpoints of the study are the following:

- **Time to treatment failure** (**TTF**): Time from the date of enrolment to discontinuation of treatment for any reason (including progression of disease, treatment toxicity, refusal and death).

- **Overall survival** (**OS**): Time from the date of enrolment until death from any cause.

- **Objective response** (**OR):** The best overall response (CR or PR) across all assessment time-points according to RECIST Criteria 1.1, during the period from enrolment to termination of trial treatment.

- **Disease control**, defined as achieving objective response or stable disease for at least 6 weeks

- **Duration of response**: Interval from the date of first documentation of objective response to the date of first documented progression or relapse.

- **Toxicity**, that is adverse events classified according to NCI CTCAE version 4.

Apart from the TTF which refers to the 'study treatment', that is combination of both drugs, time to treatment discontinuation for each drug separately (erlotinib, bevacizumab), will also be assessed. Due to the different scheduling for each drug (bevacizumab: i.v. at the 1st day of each 3-week cycle; erlotinib: p.o. daily), the following cases should be taken into consideration:

- If erlotinib is permanently discontinued (irrespectively of bevacizumab administration):

  End date for erlotininb is the date of last administration of erlotinib

- If bevacizumab is permanently discontinued (at cycle *C*), then we further identify the cases:

o   If erlotinib is also discontinued within the same cycle C:

End date for both drugs is the date of last erlotinib administration

o   If administration of erlotininb is further continued to cycle C+1:

End date for bevacizumab is day1 of cycle C+1

o   If administration of erlotininb has been discontinued in a previous cycle, before cycle C:

End date for bevacizumab is date of last administration (i.v.) of bevacizumab + 21 days (corresponding to a 3-week cycle), unless the patient experienced PD, death or withdrew from the study during cycle C (in such a case the earlier of those dates is used as end date for discontinuation of bevacizumab).

Note: In all cases, the end date of any drug cannot be later than PD, death or withdrawal date.

Another measure of interest to be investigated is relative dose intensity of each of the drugs, defined as follows:

• Relative dose intensity for erlotinib: "Totally administered dose / [150 x total duration (in days) of erlotinib treatment]" (for each patient), where 'Totally administered dose'= SUM (administered dose*days of administration).

• Relative dose intensity for bevacizumab: "Sum of actual doses for all cycles / sum of planned doses of all cycles", up to the last cycle receiving bevacizumab (for each patient).
Alternatively, relative dose intensity for bevacizumab could be defined as [Sum over all cycles of (actual dose per cycle / planned dose per cycle)] /# of cycles" i.e., *relative dose intensity for each cycle (of each patient), and then take average (for all the cycles of each patient)*

4.2 Patient accrual and baseline characteristics

A flowchart will be created to graphically depict the flow of patients and the phases of the trial. Patient accrual, overall and by T790M mutation, will be presented in tabular format for each center. In addition, cumulative accrual by T790M mutation and observed vs. expected accrual, by T790M mutation, will be graphically displayed.

With respect to baseline characteristics, first of all, a bar chart with the distribution of T790M mutation will be provided. Furthermore, patient baseline characteristics will be summarized and presented **for** all enrolled patients overall and by mutation subgroup.

All continuous variables will be summarised using the following descriptive statistics:

n (non-missing sample size), mean, 95% CI for the mean, median, maximum and minimum and Interquartile range (IQR). The frequency and percentages of observed levels will be reported for all categorical measures.

All summary tables will be structured with a column for each subgroup in the order (T790M mutated, T790M non-mutated) and a column for the total population, and will be annotated with the total population size relevant to that table/treatment, including any missing observations.

The balance of baseline characteristics between the two subgroups will be explored via the Fisher's exact test (for categorical variables) and the Mann-Whitney test for medians (for continuous variables). Barplots and histograms will be created, where appropriate, as a graphical representation of the findings.

Table 1 summarizes the baseline characteristics that are considered in the analysis. For categorical variables, the corresponding levels of each variable are also documented.

Table 1. *Summary of baseline characteristics*

Categorical characteristics
Gender
Male
Female
Smoking status
0 – Never

Table 1. *Summary of baseline characteristics*

1 – Former (≥100 cig & ≥12 months smoke-free)

2 – Current

Tumor histology

1 – Adenocarcinoma

2 – Adenocarcinoma with minor squamous cell pattern (possible

adenosquamous carcinoma)

ECOG Performance Status

0

1

2

Chemotherapy

Yes

No

Radiotherapy

Yes

No

Concomitant treatment

Yes

No

EGFR Mutations

Exon 19 (DEL19)

Exon 21 (L858R)

Localization of tumor metastasis

Adrenal

Bone

Brain

Kidney

Liver

Lung

Table 1. *Summary of baseline characteristics*

Skin
Lymph Node(s)
Other
Continuous characteristics
Age (in years)

5. EFFICACY ANALYSIS

The primary efficacy analysis will be performed in the efficacy cohort, separately in the 2 subgroups according to T790M status (substudies) as well as overall.

The goal of the primary analysis is to obtain an estimation of PFS. The overall PFS and PFS according to T790M status will be determined by the Kaplan-Meier method and compared between the two subgroups (with and without EGFR T790M mutation) by means of the logrank test.

Information on outcome (Progression free survival (PFS), time to treatment failure (TTF) (overall and by drug), and overall survival (OS)) will be presented (number of events, 12-month estimates, medians, 95% CIs), overall and within the two mutation subgroups. The log-rank test will be used to explore differences in hazard between the two mutation subgroups, while observed differences will be graphically depicted via Kaplan-Meier curves. In addition, for the T790M mutation subgroup, the number of patients (among the first 35 patients) who are progression-free at 12-months will be also separately reported.

The median follow-up time will be calculated using the reverse censoring method for OS.

Cox proportional hazard (PH) models will also be applied for PFS, OS and TTF in order to explore the significance of EGFR T790M mutation. The influence of baseline characteristics and clinical variables of interest (Table 1) will also be explored through corresponding adjusted Cox PH models. Backward elimination will be used to conclude on the statistically significant terms that should remain into the final models. The variable elimination criterion will be a p-value greater than 10%. Estimated effects (e.g. hazard ratios) with confidence intervals will be presented. Forest plots will be created as a graphical representation of the multivariate models.

Clinical efficacy will be further described by objective response rate (ORR), disease control rate (DCR) and duration of response (DR), overall and by subgroup. Regarding DR, information on the number of progressions after objective response (%) and median (95% CI) will be provided, while DR will be also graphically depicted via Kaplan-Meier curves, by subgroup. Relative dose intensity of both erlotinib and bevacizumab will be also described, overall and by subgroup. Histograms will be created, as a graphical representation of the findings. A watefall plot of best percentage change from baseline in the sum of longest tumour diameters will be also presented, by subgroup. Information on OR, will be provided as well.

The correlation of the biomarkers BRCA1 mRNA and AEG-1 mRNA and T790M with PFS and OS will be evaluated by univariate and multivariate Cox PH models. BRCA1 mRNA and AEG-1 mRNA expressions will also be appropriately discretized and their association with PFS, OS and TTF will be further evaluated through survival tables and graphically depicted through Kaplan-Meier curves. Discretization of biomarkers will be primarily based on their median values, alternative grouping, e.g. based on tertiles, could be further explored, as secondary analysis only. Any additional biomarkers will be considered as secondary analysis.

Upon data availability, the longitudinal development of EGFR mutations (including T790M) in serum will be evaluated descriptively. The relation of molecular biomarkers to EGFR TKI and bevacizumab will be explored. The feasibility of re-biopsies at the time of progression will be reported. Gene-expression arrays for decision-making for second-line treatment will be done in case of available material. The feasibility of recommending customized second-line chemotherapy based on BRCA1 and AEG-1 mRNA levels will be explored and reported.

The feasibility of extracting information on concomitant medications in a systematic way so as to correlate it with primary and secondary objectives will be investigated.

6. SAFETY ANALYSIS

The safety cohort will encompass all patients who have received at least one dose of trial treatment. Safety and tolerability of the erlotinib and bevacizumab combination will be described by tabulation of the CTCAE V4 grade and graphical representation of the corresponding distributions using bar charts.

All results will be presented overall as well as by mutation subgroup.

In particular, safety analysis will include assessment of the experience of adverse events (AE), the special cases of serious adverse events (SAE) and adverse events of special interest (AESI), the frequency of AEs (overall and separately for targeted and non-targeted AEs), the numbers of patients experiencing specific number of adverse events and their distribution by CTCAE V4 grade. Statistical tests will be performed to compare the two mutation groups.

Further exploration on AEs, with respect to their grade will include the presentation of maximum severity of AEs for each patient as well as the distribution of worst degree of severity by patient. The distribution of SAEs by provider center will also be presented.

Note: If the same adverse event is recorded more than once for a specific patient, it will be counted only once keeping the one with the higher grade.

For a closer investigation and for purposes of medical review, information on AEs will also be provided by patient. This information will consist of AE description and grade, the related cycle number, date of AE onset and end date (if the event is resolved), its potential relation to erlotinib and/or bevacizumab treatment, the subgroup each patient belongs to (T790M mutated and T790M non-mutated) and his/her specific outcome.

Summary table of all the adverse events experienced by patients with progression or death as well as the narratives of patients with SAE or AESI will also be provided.

Furthermore, the baseline symptoms experienced by some of the patients, which are not counted as adverse events –unless their grade was increased or they appeared as new events at a later period of time- will be presented separately.

Details on the definition of adverse events (as described in the protocol of the study) are available in the Appendix.

7. TECHNICAL DETAILS

Data will be analyzed using the SAS software package (version 9.3).

A second statistician, the ***reviewing statistician*** will independently reproduce all analysis and summary statistics. The ***reviewing statistician*** will have an overview of the entire analysis and will explicitly check the code producing tables and figures as well as any other pieces of code as desired.

P-values≥0.001 will be reported with two significant digits (e.g. p=0.26, p=0.026, p=0.0032); p-values less than 0.001 will be reported as "<0.001". The mean, 95% confidence limits, quantiles, and any other statistics, will be reported to one decimal. Hazard ratios (HRs) and their 95% CI's will be reported to two decimals. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 3 decimals.

8. LISTING OF TABLES AND FIGURES

Table 2 gives a tabulation of the following aspects unique to each table:

- Title

- Population

- Endpoint(s)

- Time Points or details of how to merge/summarize multiple observations

- Covariates or Subgroups used to break down summary statistics

- Which summary statistics will be calculated

- Or, what formal analysis will be used

For figures the equivalent information is summarized in Table 3 and includes the following:

- Title

- Population

- Type of figure

- Endpoint(s), and which is used for horizontal and vertical co-ordinates

- Statistic(s) used in calculating co-ordinate values used in the figure

- Covariates used within the figure used to determine colors or symbols

- Covariates used to define facets or sub-plots

Table 2. *Listing of Tables*

Table title	Population	Endpoint(s)/ Variable(s)	Time Points or how to aggregate	Covariates or Subgroups	Summary Statistics	Formal Analysis	Notes
Accrual by center by T790M mutation and overall	Efficacy cohort	Provider	NA	T790M mutation	n (%)	NA	
Summary of baseline characteristics by T790M mutation and overall	Efficacy cohort	- Gender - Smoking status - Tumor histology - ECOG performance status - Radiotherapy - Chemotherapy - Concomitant treatment - EGFR mutation Exon19/Exon21 - Localization of tumour metastasis	Baseline	T790M mutation	n (%)	Fisher's exact test	
		Age			Mean, 95% CI, median, min, max, IQR	Mann-Whitney test	
Time on follow-up (FU) and Time on treatment (TTF) by T790M mutation and overall	Efficacy cohort	FU, TTF	FU: Time from enrolment until last day of follow-up TTF: Time from enrolment until discontinuation of treatment for any reason (progression, toxicity, refusal, death)	T790M mutation	n (%), Median (Interq. Range for F-up, 95% CI for TTF)	Reverse Censoring based on OS (f-up) Kaplan – Meier analysis (both)	

Table title	Population	Endpoint(s)/ Variable(s)	Time Points or how to aggregate	Covariates or Subgroups	Summary Statistics	Formal Analysis	Notes
Objective Response (OR) by T790M mutation and overall	Efficacy cohort	OR	OR: Best overall response (CR or PR) across all assessment time-points, during the time from enrolment until termination of treatment	T790M mutation	n (%)	NA	
Disease Control (DC) by T790M mutation and overall	Efficacy cohort	DC	DC: Defined as achieving objective response (OR) or stable disease (SD) for at least 6 weeks	T790M mutation	n (%)	NA	
Duration of Response (DR) by T790M mutation and overall	Efficacy cohort responders only	DR	DR: Time from documentation of response to disease progression	T790M mutation	n (%)	NA	
Time to Bevacizumab discontinuation by T790M mutation and overall	Efficacy cohort	Time to Bevacizumab discontinuation	Time to Bev. Discont.: Time from first administration until discontinuation of Bevacizumab treatment for any reason (progression, toxicity, refusal, death)	T790M mutation	n (%) of Bevacizumab discontinuations, % on Bevacizumab at 12 months (95% C.I.), Median (95% CI)	Kaplan – Meier analysis, Log- rank test	
Time to Erlotinib discontinuation by T790M mutation and overall	Efficacy cohort	Time to Erlotinib discontinuation	Time to Erl. Discont.: Time from first administration until discontinuation of Erlotinib treatment for	T790M mutation	n (%) of Erlotinib discontinuations, % on Erlotinib at 12 months (95% C.I.), Median (95%CI)	Kaplan – Meier analysis, Log- rank test	

Table title	Population	Endpoint(s)/ Variable(s)	Time Points or how to aggregate	Covariates or Subgroups	Summary Statistics	Formal Analysis	Notes
			any reason (progression, toxicity, refusal, death)				
Descriptives for the Relative Dose Intensity (RDI) by T790M mutation and overall	Efficacy cohort	-Bevacizumab -Erlotinib	Entire period of patient X in the study	T790M mutation	Mean, 95% CI, median, min, max	Mann-Whitney test	
Progression-free Survival (PFS) according to T790M mutation and overall	Efficacy cohort	PFS	PFS: Time from enrolment until documented progression or death without documented progression or up to last day of follow-up, if event has not occurred,	T790M mutation	n (%) of PFS events, PFS% at 12 months (95% C.I.), Median (95% CI)	Kaplan – Meier analysis, Log- rank test	Analogous tables for PFS, OS, TTF will be produced by BRCA1 or AEG1 categories, EGFR subtypes and brain metastatic status in the overall cohort as well as within T790M mutation subgroups
Overall Survival (OS) according to T790M mutation and overall	Efficacy cohort	OS	OS: Time from enrolment until death, or up to last day of follow-up, if event has not occurred	T790M mutation	n (%) of deaths, OS% at 12 months (95% C.I.), Median OS (95% CI)	Kaplan – Meier analysis, Log- rank test	
Time-to-Treatment Failure (TTF) according to T790M mutation and overall	Efficacy cohort	TTF	TTF: Time from enrolment until discontinuation of treatment for any	T790M mutation	n (%) of TTF events, TTF% at 12 months (95% CI)	Kaplan – Meier analysis, Log- rank test	

Table title	Population	Endpoint(s)/ Variable(s)	Time Points or how to aggregate	Covariates or Subgroups	Summary Statistics	Formal Analysis	Notes
			reason (progression, toxicity, refusal, death)		(95% C.I.), Median TTF		
Univariate Cox proportional hazards model for PFS	Efficacy cohort	PFS	PFS: Time from enrolment until documented progression or death without documented progression or up to last day of follow-up, if event has not occurred,	T790M mutation	HR (95% CI)	Cox proportional hazards	
Multivariate Cox proportional hazards model for PFS adjusting for variables of clinical interest	Efficacy cohort	PFS	PFS: Time from enrolment until documented progression or death without documented progression or up to last day of follow-up, if event has not occurred,	T790M mutation, statistically significant baseline covariates	HR (95% CI)	Cox proportional hazards – backward elimination	Only statistically significant terms will be included
Univariate Cox proportional hazards model for OS	Efficacy cohort	OS	OS: Time from enrolment until death, or up to last day of follow-up, if event has not occurred	T790M mutation	HR (95% CI)	Cox proportional hazards	
Multivariate Cox proportional hazards model for OS adjusting for variables of clinical interest	Efficacy cohort	OS	OS: Time from enrolment until death, or up to last day of follow-up, if event has not occurred	T790M mutation, statistically significant baseline covariates	HR (95% CI)	Cox proportional hazards – backward elimination	Only statistically significant terms will be included

Table title	Population	Endpoint(s)/ Variable(s)	Time Points or how to aggregate	Covariates or Subgroups	Summary Statistics	Formal Analysis	Notes
Results from separate Cox models for PFS investigating the interaction of T790M with each of the baseline characteristics	Efficacy cohort	Baseline characteristics	At baseline	T790M mutation	n vs n, HR (95% CI)	Interaction p-value	
Adverse event overview by T790M mutation and overall	Safety cohort	AE	At baseline; at the end of each cycle; at the end of treatment visit	T790M mutation	n (%) of persons who have experienced >=1 AE	Fisher's exact test	
Serious adverse event overview by T790M mutation and overall	Safety cohort	SAE	At baseline; at the end of each cycle; at the end of treatment visit	T790M mutation	n (%) of persons who have experienced >=1 SAE	Fisher's exact test	
Adverse events of special interest overview by T790M mutation and overall	Safety cohort	AESI	At baseline; at the end of each cycle; at the end of treatment visit	T790M mutation	n (%) of persons who have experienced >=1 AESI	Fisher's exact test	
Frequency of serious adverse events by T790M mutation and overall, according to CTCAE Version 4	Safety cohort	SAE	At baseline; at the end of each cycle; at the end of treatment visit	T790M mutation	n (%)	NA	In case of multiple reports of the same SAE (grouped according to CTCAE V4) for a particular patient, the SAE is counted once.
Number of patients experiencing specific number of serious adverse events overall and by T790M mutation,	Safety cohort	SAE	At baseline; at the end of each cycle; at the end of treatment visit	T790M mutation	n	NA	

Table title	Population	Endpoint(s)/ Variable(s)	Time Points or how to aggregate	Covariates or Subgroups	Summary Statistics	Formal Analysis	Notes
according to CTCAE Version 4							
Serious adverse events overview by center	Safety cohort	SAE	At baseline; at the end of each cycle; at the end of treatment visit	Provider	n (%) of pts n(%) of events	NA	
Adverse events (AE/SAE/AESI) according to CTCAE Version 4, overall and by T790M mutation	Safety cohort	AE/SAE/AESI	At baseline; at the end of each cycle; at the end of treatment visit	T790M mutation	n (%)	NA	- In case of multiple reports of the same event for a particular patient, the AE is counted once.
Number of patients experiencing specific number of adverse events overall and by T790M mutation according to CTCAE Version 4	Safety cohort	AE/SAE/AESI	At baseline; at the end of each cycle; at the end of treatment visit	T790M mutation	n	NA	
Targeted adverse events according to CTCAE Version 4	Safety cohort	AE/SAE/AESI	At baseline; at the end of each cycle; at the end of treatment visit	T790M mutation	n (%)	NA	
Non-targeted adverse events according to CTCAE Version 4	Safety cohort	AE/SAE/AESI	At baseline; at the end of each cycle; at the end of treatment visit	T790M mutation	n (%)	NA	
Maximum severity of adverse events (AE/SAE/AESI) per patient overall and by T790M mutation	Safety cohort	AE/SAE/AESI	At baseline; at the end of each cycle; at the end of treatment visit	T790M mutation	n (%)	NA	

Table title	Population	Endpoint(s)/ Variable(s)	Time Points or how to aggregate	Covariates or Subgroups	Summary Statistics	Formal Analysis	Notes
Adverse event information by patient	Safety cohort	AE/SAE/AESI	At baseline; at the end of each cycle; at the end of treatment visit	T790M mutated	NA	NA	- If a patient experienced the same adverse event more than one times, the highest grade is presented. - If a patient discontinued first from one drug and then from the other - and since the treatment was received as a combination of the 2 drugs-, the date of a patient's discontinuation is considered to be the date he/she discontinued from the first one.
Narratives of patients with a SAE and/or an AESI	Safety cohort	SAE/AESI	NA	NA	NA	NA	

Table 3. *Listing of Figures*

Title	Population	Type of graph	Horizontal Variables	Vertical Variables	Groupings	Statistics	Facets
Patients enrolment	All patients registred in the database	Flowchart	NA	NA	NA	NA	NA
Cumulative Overall Accrual by T790M mutation	Efficacy cohort	Line graph	time	frequency	NA	NA	NA
Expected vs. observed accrual by T790M mutation	Efficacy cohort	Line graph	time	frequency	T790M mutation	NA	NA
Distribution of T790M mutation	Efficacy cohort	Bar chart	percentage	percentage	T790M mutation	percentage	NA
Time on Follow-up by T790M mutation	Efficacy cohort	KM	time	probability	T790M mutation	Survival estimates	Subgroups and overall
Waterfall plot of best percentage change from baseline in the sum of target lesion diameters	Efficacy cohort	Waterfall plot	patient	Change from baseline	PD, SD, PR, CR	NA	NA
Duration of Response by T790M mutation for patients who reported objective response	Efficacy cohort (responders)	KM	time	probability	T790M mutation	Survival estimates	Subgroups and overall
Time to discontinuation from Bevacizumab by T790M mutation	Efficacy cohort	KM	time	probability	T790M mutation	Survival estimates	Subgroups and overall
Time to discontinuation from Erlotinib by T790M mutation	Efficacy cohort	KM	time	probability	T790M mutation	Survival estimates	Subgroups and overall
Relative Dose Intensity (RDI) of Bevacizumab by T790M mutation	Efficacy cohort	Histogram	Relative Dose Intensity	percentage	T790M mutation	Mean, median, range, skewness, variance	NA
Relative Dose Intensity (RDI) of Erlotinib by T790M mutation	Efficacy cohort	Histogram	Relative Dose Intensity	percentage	T790M mutation	Mean, median, range, skewness, variance	NA

Title	Population	Type of graph	Horizontal Variables	Vertical Variables	Groupings	Statistics	Facets
Progression-free Survival by T790M mutation$	Efficacy cohort	KM	time	probability	T790M mutation	Survival estimates	Subgroups and overall
Overall Survival by T790M mutation$	Efficacy cohort	KM	time	probability	T790M mutation	Survival estimates	Subgroups and overall
Time on treatment by T790M mutation$	Efficacy cohort	KM	time	probability	T790M mutation	Survival estimates	Subgroups and overall

$ Analogous KM plots for PFS, OS, TTF will be produced by BRCA1 or AEG1 categories, EGFR subtypes and brain metastatic status in the overall cohort as well as within T790M mutation subgroups

REFERENCES

TR Fleming (1982). One-sample multiple testing procedure for Phase II clinical trials. Biometrics, 38: 143-151.

R Rosell, MA Molina, C Costa, et al (2011). Pretreatment EGFR T790M Mutation and BRCA1 mRNA Expression in Erlotinib-Treated Advanced Non-Small-Cell Lung Cancer Patients with EGFR Mutations. Clin Cancer Res, 17: 1160-1168.

R Simon (1989). Optimal two-stage designs for Phase II clinical trials. Controlled Clinical Trials, 10: 1-10.

APPENDIX

IMPORTANT INFORMATION AS DESCRIBED IN THE PROTOCOL

A.1 STUDY OBJECTIVES AND ENDPOINTS

A.1.1   Study Objectives

Primary Objective

To determine progression-free survival (PFS) of patients with advanced non-squamous NSCLC harboring at diagnosis EGFR mutations with and without T790M mutation treated with the combination of erlotinib and bevacizumab.

Secondary Objectives

i.   To evaluate secondary measures of clinical efficacy including overall survival (OS), time to treatment failure (TTF), objective response rate (ORR), disease control rate (DCR) and duration of response (DR).

ii.   To assess the safety and tolerability of the erlotinib and bevacizumab combination.

iii.   To evaluate the correlation of BRCA1 mRNA and AEG-1 mRNA expression and T790M with progression-free survival.

iv.   To monitor EGFR mutations (including T790M) in serum and plasma longitudinally.

v.   To evaluate molecular biomarkers related to EGFR TK1 and bevacizumab.

vi.   To determine the feasibility of re-biopsies at the time of progression and gene-expression arrays for decision-making for second-line treatment.

vii.   To study the feasibility of recommending customized second-line chemotherapy based on BRCA-1 and AEG-1 mRNA levels.

A.1.2   Endpoints

Primary Endpoint

The primary endpoint of the study is Progression-free Survival (PFS), defined as the time from the date of enrolment until documented progression or death, whichever occurs first.

Secondary Endpoints

Secondary endpoints include Overall Survival (OS), Time to treatment failure (TTF), Time to treatment discontinuation (TTD), Objective Response (OR), Disease Control (DC), Duration of response (DR) and adverse events graded according to CTCAE V4.0. More specifically OS is defined as the time from the date of enrolment until death from any cause. In addition, TTF is defined as time from the date of enrolment to discontinuation of treatment for any reason, including progression of disease, treatment toxicity, refusal and death. As for TTD, two different definitions of treatment discontinuation are considered: i) all cause treatment discontinuation, and ii) discontinuation of treatment for patient decision. OR is defined as the best overall response (CR or PR) across all assessment time-points according to RECIST Criteria 1.1, during the period from enrolment to termination of trial treatment. DC is defined as achieving objective response or stable disease for at least 6 weeks. Finally, DR is defined as the interval from the date of first documentation of objective response by RECIST to the date of first documented progression or relapse.

A.1.3   Statistical Hypotheses

The specific statistical hypotheses that are to be tested in order to assess the primary objective of the study are listed below. A one-sided significance level of 0.05 will be used for statistical testing.

- **Substudy 1: Patients with EGFR T790M mutation**

    H_0 :   The 12-months PFS ≤ 40%

    H_1 :   The 12-months PFS ≥ 63%

- **Substudy 2: Patients without EGFR T790M mutation**

    H_0 :   The 12-months PFS ≤ 50%

    H_1 :   The 12-months PFS ≥ 65%

A.2 STUDY METHODS

A.2.1   General Study Design and Plan

This is a multinational, multi-center trial of erlotinib plus bevacizumab in patients with advanced non-squamous NSCLC harbouring EGFR mutations, including two separate phase II substudies. Only EGFR mutated patients may enter the study. Approximately 1135 patients with advanced non-small cell lung cancer are expected to be screened to include 102 patients with activating EGFR mutations. Patients may continue treatment on this trial as long as there is evidence of clinical benefit in the judgment of the investigator. Trial subjects will receive trial medication up to 18 months after the inclusion of the last patient. All patients who are still benefiting from their treatment at that time have to be switched to commercial drug which will be reimbursed by Roche.

Patient accrual is expected to be completed within 12 months including a run-in-period of 3 months. Treatment and follow-up is expected to extend the study duration to a total of 48 months. Patients will be followed until death – thus follow-up estimated up to 3 years following the enrolment of the last patient.

The trial will end with the preparation of the final report, scheduled for month 54 after the inclusion of the first patient.

A.2.2   Patient Selection

Only patients with a centrally confirmed status of EGFR mutation are eligible. The mutation status may first be assessed locally by a certified laboratory, but for all patients tumor material has to be submitted to the reference laboratory (BELIEF Protocol, Section 9) for immediate confirmatory testing. If the EGFR mutation status is not confirmed by the central reference laboratory, the patient will be taken off trial and will undergo an end-of-treatment assessment. Such a patient will be followed up per protocol. Written informed consent needs to be obtained prior to shipment of tissue to the central laboratory.

Patients should only be selected and consented for screening if they fulfill the inclusion and exclusion criteria described in Sections 7.1 and 7.2 of the BELIEF Protocol respectively.

A.2.3   Patient Screening and Registration

This trial will use a web-based registration system. Each participating centre will access the registration system directly. Specific details for registration of patients are in the **"BELIEF Procedures Manual**" which will be available on the ETOP website (www.etop-eu.org).

A.2.4   Early Stopping Rules

Substudy 1 contains an early look at the end of the first stage, when 8 patients have reached 12 months follow-up. If 3 or less patients reach 12 months without a progression-defining event, then the results of the substudy will be reported immediately and the Steering Committee will decide whether the patients still on treatment should stop it. Otherwise the substudy will continue as planned until the final evaluation.

A.2.5   Study Variables

Trial schedule of events

	Scree-ning [1]	≤ 28 days prior enrol-ment	At enrol-ment	Cycle 1	Cycles 2, 4, ... (even cycles)	Cycles 3, 5, ... (odd cycles)	At pro-gression	End of treat-ment visit [4]	Follow-up: every 12 weeks
Days			0	1	22,64,...	43,85,...			
Weeks			0	1	4,10,...	7,13,...			
Written informed consent [2]	X								
Medical history	X								
Tumour material [1]	X								
Blood sample [4]			X			X [4]	X		
Physical exam, PS, blood pressure, weight		X		X	X	X		X	X
Baseline symptoms		X							
Haematology		X		X	X	X		X	
Renal function: serum creatinine, urine dipstick		X		X	X	X		X	
Hepatic function: ALT, AST, AP, Bilirubin		X		X	X	X		X	
Coagulation: INR		X							
Pregnancy test if applicable		X							
Tumour assessment [3]		X				X	X	X [5]	X
Adverse events				X	X	X		X	
Bevacizumab				X	X	X			
Erlotinib (daily)				X	X	X			
Second-line therapy									X

1 – screening must be done within 28 days before treatment start; tumour material from biopsy and/or original surgery

2 – before any trial-specific intervention

3 – CT of thorax and abdomen prior to start of cycles 3, 5, 7, 10, 13, 16, 20, 24, 28; in case of treatment discontinuation without progression weeks 6, 12, 18, 27, 36, 45, 57, 69, 81 and every 12 weeks until progression. CT scan of brain is not mandatory and only recommended in case of clinically suspected brain metastasis.

4 – A serum sample for translational research should be taken at baseline, at the time of first documentation of response, and at the time of progression.

5 – For patients who discontinued study treatment due to toxicity rather than progressive disease, restaging (CT scan chest and abdomen, brain scan if applicable) should be repeated if not already performed within 30 days prior to the last dose of study treatment

Baseline evaluations (within 28 days prior to registration)

- Medical history including symptoms, smoking history, medications, comorbidities and allergies

- Physical examination including blood pressure [mmHg], ECOG performance status (see definition in **BELIEF procedures manual)**, and body weight [kg]

- Haematology: haemoglobin, leukocytes, platelets

- Renal function: serum creatinine and creatinine clearance calculated according to Cockroft-Gault, urine dipstick for proteinuria

- Hepatic function: ALT, AST, AP, Bilirubin

- Coagulation: INR

- Pregnancy test for women with childbearing potential

- CT scan of thorax and upper abdomen with i.v. contrast (alone or in combination with PET) to determine measurable disease according to RECIST v1.1 (at least one lesion outside of irradiated areas that can be measured in at least one dimension as ≥ 10 mm, or ≥ 15 mm in case of lymph nodes). In the presence of clinically suspected metastases outside the thorax (e.g. brain, bone or lower abdomen), additional CT of the affected body part is recommended.

- CT scan of brain is not mandatory and only recommended in case of clinically suspected brain metastasis

Before enrolment

- For central confirmation of EGFR exon 19 deletion (del19) or exon 21 mutation (L858R), send to central laboratory:

  - 1 block of FFPE tissue from biopsy, surgery or cytology
  - Pathology report
  - Local EGFR mutation report (if applicable)

Before start of treatment

- Take a blood sample for routine evaluations plus a serum and plasma sample for translational research

Routine evaluations before and during trial treatment

On day 1 of every 3-week treatment cycle:

- Recording of symptoms / adverse events

- Physical examination including blood pressure, performance status, and body weight

- Haematology: haemoglobin, neutrophils, platelets

- Serum creatinine

- Hepatic function: ALT, AST, AP, Bilirubin

- Urine dipstick

Trial – specific evaluations during treatment

These evaluations will occur more frequently during the first cycles, and be phased out especially for patients who stay on treatment for a prolonged period. They should be performed within 72h before the start of the subsequent cycle:

- Prior to the start of cycles 3, 5, 7, 10, 13, 16, 20, 24, 28 and every 4 cycles until progression (weeks 6, 12, 18, 27, 36, 45, 57, 69, 81; and then every 12 weeks)

- CT thorax and upper abdomen

- Blood serum and plasma sample for translational research at the first time of documented response

Evaluation in the follow-up phase before progression

Patients who discontinue treatment before progression should have the following assessments at the same timepoints as patients still on treatment (weeks 6, 12, 18, 27, 36, 45, 57, 69, 81; and every 12 weeks until progression):

- Physical examination

- CT thorax and upper abdomen (plus further imaging if applicable)

- Documentation of further treatments

Evaluation at progression

Each patient will receive trial treatment until documented progression. At progression, do the following:

- CT thorax and upper abdomen, document progression on the respective CRF

- Blood serum and plasma sample for translational research

- Tumor re-biopsy should be considered for testing of further predictive markers and translational research at the central reference laboratory (Pangaea Biotech)

End of treatment visit

At the end of the trial treatment and **irrespective of the reason for stopping treatment**, a post treatment visit at the center is to be scheduled after 30 (to 40) days following last treatment day. The following procedures should be performed:

- Recording of symptoms
- Physical examination including blood pressure
- Haematology: haemoglobin, neutrophils, platelets
- Hepatic function: ALT, AST, AP, Bilirubin
- Serum creatinine
- Urine dipstick
- CT thorax and upper abdomen, if not done within the last 30 days

Evaluation after progression

Patients with progression will end trial treatment and should have documented

- survival and
- further lines of treatment

up to 3 years after the inclusion of the last patient.

A.2.6   Primary Comparisons

The goal of the primary analysis of the study is to obtain an estimation of PFS. All eligible and treated patients will be included in this analysis. The overall PFS and PFS according to T790M status will be determined by the Kaplan-Meier method and compared between the two subgroups (with and without EGFR T790M mutation) by means of the log-rank test.

A.3 SAMPLE SIZE

The EAST software package is used for sample size calculations (EAST 5, Version 5.4.0.0, Cytel Inc. 2010).

The sample size calculations are based on the percentage of patients with EGFR T790M mutation and corresponding median PFS by EGFR T790M mutation under erlotinib treatment, as presented in the recently published results by Rosell et al. (2011).

The trial encompasses two phase II substudies run in parallel, and requires a total of 102 patients:

- *Substudy 1*: 35 patients with EGFR T790M mutation
- *Substudy 2*: 67 patients without EGFR T790M mutation

The first substudy follows Simon's two-stage design and a decision can be reached at the first stage, while the second substudy follows Fleming's single stage design.

Substudy 1: Patients with EGFR T790M mutation

The primary objective is the estimation of PFS in patients with EGFR T790M mutation, treated with bevacizumab and erlotinib. The median PFS is expected to improve to 18 months. A 12-month PFS of 40%, corresponding to a median of 9 months, will be considered inadequate (since a 12 month median PFS is achieved by erlotinib alone), while the target value will be a 12-month PFS of 63% (corresponding to a median PFS of 18 months, similar to the one observed with erlotinib alone in the subgroup of patients without EGFR T790M mutation).

Simon's optimal two-stage design (Simon, 1989), with a significance level of 5% and power of 80%, will be used to test the null hypothesis that the 12 month PFS≤40% versus the alternative that PFS≥63%. If in the first stage from a total of 8 patients, 4 or more patients reach 12 months without a progression-defining event, then the substudy will proceed so as to evaluate the 12-month PFS for a total of 35 patients. If 19 or more patients reach 12 months without a PFS event, treatment with the combination of erlotinib and bevacizumab will warrant further study in a Phase III trial in the subgroup of patients with EGFR T790M mutation.

It is anticipated that the decision on the first stage of substudy 1 will not be reached before the study is fully accrued, since 12 months of follow-up will be required for at least 8 patients with EGFR T790M mutation, while the anticipated full study accrual will be completed within 12 months.

In case, the EGFR T790M mutation is identified on only 30% of the patients (instead of the assumed 35%), the decision that can be reached at the end of the study with significance level of 5% and power of 80% will correspond to an alternative of PFS≧66% or a median PFS≧20 months.

Substudy 2: Patients without EGFR T790M mutation

The primary objective is the estimation of PFS in patients with EGFR mutation without T790M mutation, treated with bevacizumab and erlotinib. The median PFS is expected to be approximately 18 months or higher. A 12-month PFS of 50% will be considered inadequate (since 50% is achieved by erlotinib alone in the subgroup of patients with EGFR T790M mutation), while the target value will be a 12-month PFS of 65% (corresponding to a median PFS of 19 months, similar to the one observed with erlotinib alone in the subgroup of patients without EGFR T790M mutation).

A sample size of 67 patients is required by Fleming's single stage design (Fleming, 1982), with significance level of 5% and power of 80%, to test the null hypothesis that the 12-month PFS≤50% versus the alternative that PFS≥65%.

A.4 ADVERSE EVENTS

A.4.1 Adverse Events (AE)

The main criterion for tolerability is the occurrence of toxicities and adverse events. The severity and causality will be classified according to the NCI CTCAE Version 4. The CTCAE is available for downloading on the internet, see Appendix 3 (http://evs.nci.nih.gov/ftp1/CTCAE/About.html).

An adverse event is defined as any untoward medical occurrence that occurs from the first dose of study medication until 30 days after the final dose, regardless of whether it is considered related to a medication. In addition, any known untoward event that occurs subsequent to the adverse event reporting period that the investigator assesses as possibly related to the protocol treatment should be considered an adverse event. Symptoms of the targeted cancer (if applicable) should not be reported as adverse events.

A.4.2   Severity Grade

The adverse event severity grade provides a qualitative assessment of the extent or intensity of an adverse event, as determined by the investigator or as reported by the subject. The severity grade does not reflect the clinical seriousness of the event, only the degree or extent of the affliction or occurrence (e.g. severe nausea, mild seizure), and does not reflect the relationship to study drug.

Severity grade for other adverse events not covered in the toxicity grading scale:

1 =Grade 1	Mild
2 =Grade 2	Moderate
3 =Grade 3	Severe
4 =Grade 4	Life- threatening
5 =Grade 5	Fatal

A.4.3   Serious Adverse Events (SAE)

SAEs during trial treatment

An SAE is defined in general as any undesirable medical occurrence/adverse drug experience that occurs during or within 30 days after stopping study treatment that, at any dose, results in any of the following:

- is fatal (any cause)
- life-threatening,
- requires or prolongs inpatient hospitalization,

- results in persistent or significant disability/incapacity
- is a congenital anomaly or birth defect
- is a secondary malignancy
- requires significant medical intervention

Second (non-NSCLC) malignancies are always considered SAEs, no matter when they are diagnosed. These events should be reported on the Serious Adverse Event Forms.

Other significant/important medical events which may jeopardize the patient are also considered serious adverse events. Serious also includes any other event that the investigator or the ETOP Safety Office judges to be serious or which is defined as serious by the regulatory agency in the country in which the event occurred.

An unexpected adverse event is one that is not listed as a known toxicity of the investigational drug in the summary of product characteristics.

A related adverse event is one for which the investigator assesses that there is a reasonable possibility that the event is related to the investigational drug. All adverse events judged as having a reasonable suspected causal relationship to the trial medication qualify as adverse reactions.

Events not considered to be serious adverse events are hospitalizations occurring under the following circumstances:
- elective surgery;
- occur on an outpatient basis and do not result in admission (hospitalization < 24h);
- are part of the normal treatment or monitoring of the studied treatment;
- progression of disease.

SAEs after end of trial treatment
During the follow-up phase (starting 30 days after end of trial treatment), the following events have to be reported as SAE:
- fatal, life-threatening and other medically significant events possibly, probably or definitely related to late effects of trial treatment

- disabling events

- second primary cancer

- congenital anomaly

- pregnancy

In the case of pregnancy occurring during the course of the trial or within 1 year after treatment discontinuation, the investigator shall immediately notify this by completing the pregnancy reporting form. The investigator shall ensure that the case is followed up to the end of the pregnancy and supply a final report on the outcome.

A.4.4   Adverse Events of Special Interest (AESI)

The following adverse events are of special interest and will be documented if observed:

- Interstitial lung disease (ILD)-like events

- Haemorrhage, with a focus on haemoptysis and CNS bleeding

- Hypertension

- Proteinuria

- Arterial/venous thromboembolism events

- Wound healing complications

- Congestive heart failure

- Fistulae

- Gastrointestinal perforations

- Reversible posterior leucoencephalopathy syndrome