Mylan GmbH

MYL-1402O-3001

Multicenter, Double-Blind, Randomized, Parallel-Group Study to Assess the Efficacy and Safety of MYL-1402O Compared with Avastin®, in the First-Line Treatment of Patients with Stage IV Non-Squamous Non-Small Cell Lung Cancer

21Aug2019

Statistical Analysis Plan

Version 2.0

Prepared by:

PPD

3900 Paramount Parkway

Morrisville, NC 27560

Prepared for:

Mylan GmbH

Thurgauerstrasse, 40

CH-8050 Zürich, Switzerland

Signature Page

Statistical Analysis Plan: MYL-1402O-3001

Multicenter, Double-Blind, Randomized, Parallel-Group Study to Assess the Efficacy and Safety of MYL-1402O Compared with Avastin®, in the First-Line Treatment of Patients with Stage IV Non-Squamous Non-Small Cell Lung Cancer

Jiali Tang, Principal Biostatistician

Jiali Tang
Principal Biostatistician
I approved this document
21 Aug 2019 17:20:13 +00:00
DocuSign

Signature
Trial Biostatistician
PPD Inc.
3900 Paramount Parkway
Morrisville, NC 27560, USA

Date of Signature

Qiuhong Zhang, Principal Biostatistician

Qiuhong Zhang
Principal Biostatistician II
I approve this document
21 Aug 2019 18:44:30 -04:00
DocuSign

Signature
Senior Biostatistics Reviewer
PPD Inc.
3900 Paramount Parkway
Morrisville, NC 27560, USA

Date of Signature

Katrin Beckmann, Principal Statistician

K. Beckmann

Signature
Mylan Healthcare GmbH
Hannover, Germany

22 Aug 2019 08:40 CEST

Date of Signature

TABLE OF CONTENTS

List of Abbreviations

ADA	antidrug antibody
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophils count
aPTT	activated partial thromboplastin time
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BDR	blinded data review
BMI	body mass index
BOR	best overall response
BSA	body surface area
BSSR	blinded sample size re-estimation
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CP	carboplatin and paclitaxel
CR	complete response
eCRF	electronic Case Report Form
CSR	Clinical Study Report
CT	computerized tomography
DCR	disease control rate
DOR	duration of response
DSMB	data safety monitoring board
ECOG	Eastern Cooperative Oncology Group
EGFR	epidermal growth factor receptor
EML4-ALK	echinoderm microtubule-associated protein-like 4-anaplastic lymphoma kinase
EOT	end of treatment
eTMF	electronic trial master file
FDA	United States Food and Drug Administration

GCP	Good Clinical Practice
HB	Hemoglobin
HR	hazard ratio
ICH	International Conference on Harmonisation
IMP	investigational medicinal product
ITT	intent-to-treat
IWRS	interactive web response system
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified intent-to-treat
MRI	magnetic resonance imaging
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
nsNSCLC	non-squamous non-small cell lung cancer
ORR	objective response rate
OS	overall survival
PD	progressive disease
PFS	progression-free survival
PK	Pharmacokinetic
PLT	Platelet
PopPK	population PK
PP	per protocol
PR	partial response
PT	preferred term
PVG	Pharmacovigilance
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SOC	system organ class
SS	safety set
TBILI	total bilirubin

TEAE	treatment-emergent adverse events
ULN	upper limit of normal
WHO	World Health Organization

1. Revision History

Version	Date	Changes from Previous Version
0.1	19Aug2016	Initial version
0.2	10Feb2017	Updates per Mylan comments
1.0	10May2018	Final approved version
2.0	21Aug2019	Updated per Protocol Amendment v3.0

2. Introduction

This document describes the planned statistical analyses for Protocol MYL-1402O-3001 (v3.0, 19 February 2019). If there are any deviations from this statistical analysis plan (SAP), they will be detailed in a note-to-file and stored in the Electronic Trial Master File (eTMF) and also described in the Clinical Study Report (CSR).

3. Study Objectives

3.1. Primary Objective

The primary objective of this study is to compare the objective response rate (ORR) of MYL-1402O with that of Avastin®, in combination with carboplatin and paclitaxel (CP) chemotherapy during the first 18 weeks of first-line treatment in patients diagnosed with stage IV non-squamous non-small cell lung cancer (nsNSCLC).

3.2. Secondary Objectives

The secondary objectives of this study are as follows:

- Assess the safety profile of MYL-1402O as compared with that of Avastin® when administered in combination with CP as first-line treatment for stage IV nsNSCLC and when administered alone in the maintenance setting.
- Assess other efficacy parameters at Week 18 and 42: disease control rate (DCR), duration of response (DOR), progression-free survival (PFS), and overall survival (OS) rate of

MYL-1402O as compared to Avastin® when administered in combination with CP as first-line treatment for stage IV nsNSCLC.

- Assess the potential immunogenicity during 42 weeks of treatment of MYL-1402O as compared with that of Avastin®.
- Compare the pharmacokinetic (PK) profile of MYL-1402O and Avastin® using a population PK (Pop PK) approach.

4. Investigational Plan

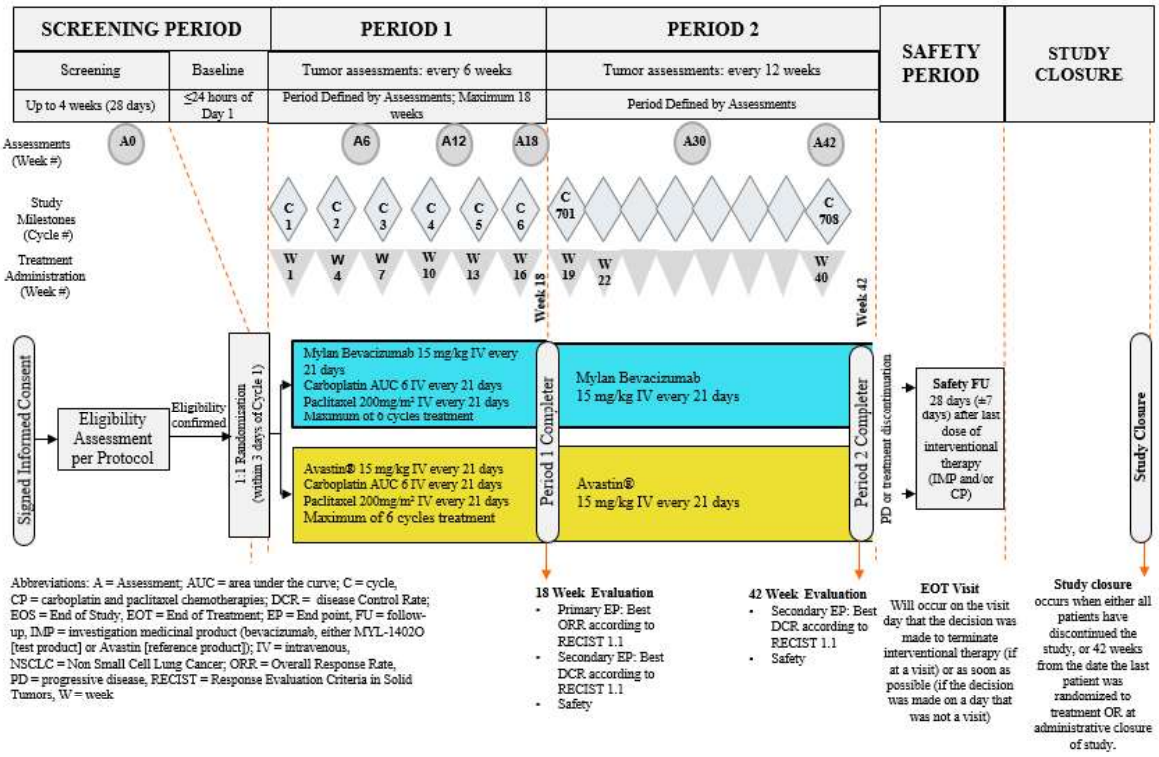
4.1. Overall Study Design and Plan

The study is a multicenter, randomized, double-blind, 2-arm, parallel-group, equivalence study to evaluate patients with Stage IV nsNSCLC when treated in first-line with bevacizumab (either MYL-1402O [test product] or Avastin® [reference product]) in combination with CP and subsequently, with monotherapy of bevacizumab (either MYL-1402O or Avastin®).

The study is comprised of 4 periods: Screening/Baseline Period, Period 1, Period 2, and Extended Treatment Period. For Period 1, the length of the treatment period is 18 weeks and it is the period where the primary endpoint will be assessed. During Period 1 tumor assessments are done every 6 weeks. Period 2 is a 24 week period where tumor assessments are done every 12 weeks.

The periods are shown in Figure 4.1-1 and are described in detail below. Assessments will be conducted per the Schedule of Events in Clinical Study Protocol.

Figure 4.1-1 Study Design



The maximum planned study participation per patient is approximately 47 weeks.

Screening/Baseline Period: In this period, which has a duration of up to 4 weeks (or ≤28 days prior to first dose [Day 1 of Cycle 1]), tumor imaging assessments must be performed no more than 28 days prior to first dose (Day 1 of Cycle 1). The time between randomization (≤3 days prior to Day 1 of Cycle1) and the first dose (Day 1 of Cycle 1) must be included in the 28-day window for tumor imaging assessment. Patients will be randomly assigned in a 1:1 ratio to one of the two treatment arms (MYL-1402O or Avastin®) within 3 days before the start of Period 1 (first dose of bevacizumab plus CP on Day 1 of Cycle 1). Baseline assessments must be performed within 24 hours of Day 1 of Cycle 1 to confirm that the patient has remained eligible for treatment.

Period 1: Patients will begin Period 1 receiving bevacizumab combination therapy (MYL1402O or Avastin®, plus CP) on Day 1 of Cycle 1 for up to 6 cycles of therapy. Each cycle will consist of 3 weeks (21 days ± 3 days) and a cycle will start with the administration of bevacizumab (as either MYL-1402O or Avastin®). Investigational medicinal product

(IMP) includes either MYL1402O or Avastin®; interventional therapy includes MYL1402O, Avastin®, and CP.

In this period, as indicated in protocol, tumor assessments should be performed every 6 weeks after Day 1 of Cycle 1 (the first dose of bevacizumab, either MYL-1402O or Avastin®) and continuing every 6 weeks through Week 18.

Tumor assessments must be done consistently throughout the study, using a computerized tomography (CT) scan of thorax and abdomen and using the same technique - CT or magnetic resonance imaging (MRI), slice interval, contrast agent. Period 1 will include 3 assessments at the pre-specified time points regardless of delays in treatment cycles.

Period 2: A patient is eligible to continue into Period 2 if he/she has a stable disease or better response (i.e., complete response [CR] or partial response [PR]) by RECIST 1.1 criteria at the Period 1 Week 18 tumor assessment. Eligible patients will continue to receive bevacizumab (either MYL-1402O or Avastin®) every 3 weeks as monotherapy.

In this period, tumor assessments will occur every 12 weeks (+/- 3 days) after the end of Period 1 (Week 18), at pre-specified time points (i.e., Week 30 and Week 42), regardless of delays of the cycles of treatment, until progressive disease (PD) or discontinuation of interventional therapy (for any reason), or withdrawal of consent. The Week 42 efficacy assessments will be used to evaluate the secondary endpoints. Patients will be assessed for tumor response using the same type of radiographic examinations used in Period 1 (e.g., CT or MRI of the thorax and abdomen, contrast agent, slice intervals) for ongoing evaluation of tumors.

If a patient is discontinued from treatment prior to completion of Period 2 due to an Adverse Event, tumor assessment will be performed as per schedule until week 42, patient is started on a new anticancer treatment (surgery, radiotherapy or systemic treatment) or PD, and will be followed for OS till week 42.

Extended Treatment Period: During the Extended Treatment Period, continued bevacizumab monotherapy may be provided to any patient that has maintained stable disease or better response (CR or PR) by RECIST 1.1 criteria at Week 42.

During the Extended Treatment Period, a patient will receive bevacizumab monotherapy (either MYL-1402O or Avastin®) until PD or discontinuation of treatment or termination of study. Safety assessments will be made for each cycle of therapy.

Safety Follow-Up Visit: The Safety Follow-Up Visit is designed to capture safety assessments for any patient(s) that completed Period 2 or discontinued the IMP at any time during the study for any reason (e.g., PD, treatment intolerance to bevacizumab). Safety follow-up visit will occur at 28 days (±7 days) after the last dose of bevacizumab during an office visit. The End of Treatment (EOT) Visit should occur after the last dose of any interventional therapy, mainly to capture the reason for discontinuation of the first-line treatment. The EOT Visit may occur at the same time as the safety follow-up visit.

Study Closure: Study closure will occur when one of the following is reached:

1. All patients have discontinued the study
2. 42 weeks from the date the last patient was randomized to treatment
3. Administrative closure of study.

4.2. Study Endpoints

4.2.1. Primary Endpoints

The primary efficacy endpoint is the ORR as assessed by independent reviewers during the first 18 weeks, assessed according to RECIST1.1.

4.2.2. Secondary Endpoints

The secondary efficacy endpoints will consist of the following:

- DCR (best overall response = CR, PR, or stable disease) during the first 18 and 42 weeks

- PFS, defined as the time from randomization to the first documentation of PD or to death due to any cause, whichever comes first; PFS rate will be calculated at 18 and 42 weeks, median PFS will be determined at 42 weeks.
- OS, defined as the time from randomization to date of death due to any cause, OS rates will be calculated at 18 and 42 weeks.
- DOR, is defined as the time from start of the first documentation of objective tumor response (CR or PR) to the first documentation of tumor progression (i.e., PD) or to death due to any cause, whichever comes first.

4.2.3. Pharmacokinetic Endpoints

The pharmacokinetics endpoints will be as follows:

- Pop PK measures of exposure of MYL-1402O and the reference product Avastin® (e.g., AUC, Cmax, Cmin, CL, Vc, and the terminal elimination half-life).

4.2.4. Safety Endpoints

The safety endpoints will be as follows:

- Incidence, nature, and severity of AEs including adverse drug reactions graded according to CTCAE.
- Detection of antibodies to bevacizumab.

5. Overview of Planned Analyses

5.1. Interim Analysis

A blinded sample size re-estimation (BSSR) interim analysis will be conducted when at least 30% of the required patients have either discontinued early or completed Week 18 on study. This interim analysis will be used to estimate the ORR at Week 18 for the total sample in the intent-to-treat (ITT) and modified ITT sets (See Definition in Section 6.5). The ITT set will be the primary set for the BSSR interim analysis. The data cut-off date will be the date

when the last patient in the analysis population completes the Week 18 assessment. Three outcomes for the estimated overall ORR are possible

- The ORR in the interim analysis <= 36 %

Then sample size will be re-estimated for the risk ratio and increased up to a maximum of 670 patient to maintain the power for testing for FDA requirement.

- The ORR in the interim analysis is > 36 % and < 40%

Then sample size will be unchanged.

- The ORR in the interim analysis >= 40%

Then sample size will be re-estimated for the risk difference and increased up to a maximum of 670 patients to maintain the power for testing EMA requirement.

The results of the interim analysis will include the overall ORR in the analysis sets and breakdown of the ORR by all prognostic factors such as, age (<65 vs. >=65 years), gender (male vs. female), race (White vs. Non-White, Asian vs. Non-Asian), smoking status (smoker vs. non-smoker), number of metastasis sites (one vs. multiple), prior radiation therapy (yes vs. no), prior adjuvant chemotherapy (yes vs. no), baseline ECOG (0 vs. 1), EGFR mutation status (unknown vs. negative), EML4-ALK alternations (unknown vs. negative), region (Europe, Asia, Rest of World). Furthermore descriptive statistics will be provided for demographics as specified in Section 8.1, patient disposition as specified in Section 7.1. ORR will be calculated as the proportion of patients with confirmed CR or PR based on the number of patients included for the analysis. Patients with response that cannot be evaluated per RECIST1.1 or without any tumor assessment prior to or on Week 18 will be scored as non-responders. If the ORR is estimated to be equal to or less than 36%, then the required sample size to meet FDA requirements will be adjusted to maintain the power of the study at 80% (Table 5.1-1 based on Farrington & Manning's likelihood score test). If the ORR is estimated to be equal to or greater than 40%, then the required sample size to meet EMA requirements will be adjusted to maintain the power of the study at 80% (Table 5.1-2 based on Farrington & Manning's likelihood score test).

Table 5.1-1 Sample Size vs. ORR based on Test of Ratio of Proportion

Power	Sample Size/Arm	ORR (%)	Target Alpha	Sample Size
0.8010	335	35.0	0.05	670
0.8008	332	35.2	0.05	664
0.8007	329	35.4	0.05	658
0.8004	326	35.6	0.05	652
0.8001	323	35.8	0.05	646
0.8013	321	36.0	0.05	641

Table 5.1-1 Sample Size vs. ORR based on Test of Difference of Proportion

Power	Sample Size/Arm	ORR (%)	Target Alpha	Sample Size
0.8016	334	50.0	0.025	668
0.8001	333	49.0	0.025	666
0.8008	333	48.0	0.025	666
0.8002	332	47.0	0.025	664
0.8000	331	46.0	0.025	662
0.8003	330	45.0	0.025	660
0.8010	329	44.0	0.025	658
0.8005	327	43.0	0.025	654
0.8004	325	42.0	0.025	650
0.8007	323	41.0	0.025	646
0.8016	321	40.0	0.025	642

The BSSR will be conducted by a statistician independent from Mylan (e.g., the CRO blinded statistician). This person will remain blinded to the true randomization that each patient received. Results including a recommendation for the re-estimated sample size will

be shared with Mylan project statistician. Mylan will make the decision for the sample adjustment based on the results.

Simulations for estimating the actual type I error rate have been performed for the primary endpoints of risk ratio and the risk difference for a study conduct with and without sample size increase. The results of 10000 simulation are displayed in the graphs below:

Figure 5.1-1 Power versus ORR (With and Without BSSR) – Difference

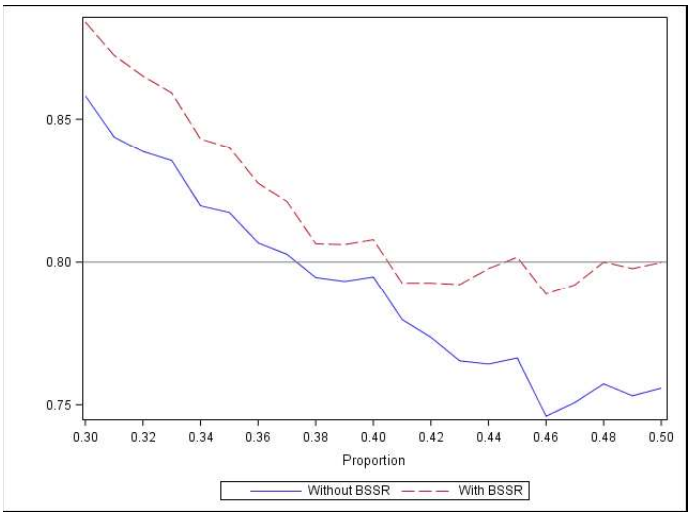
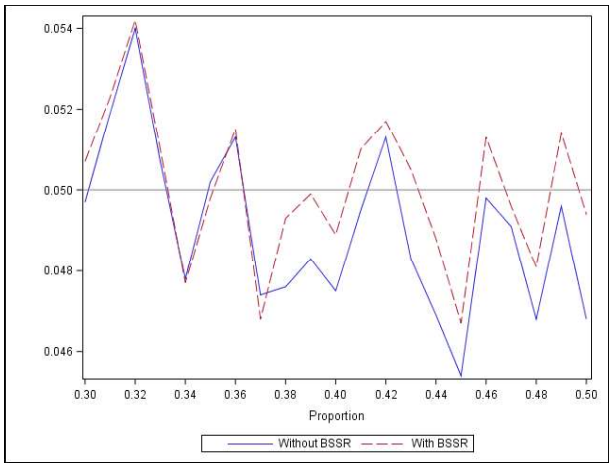
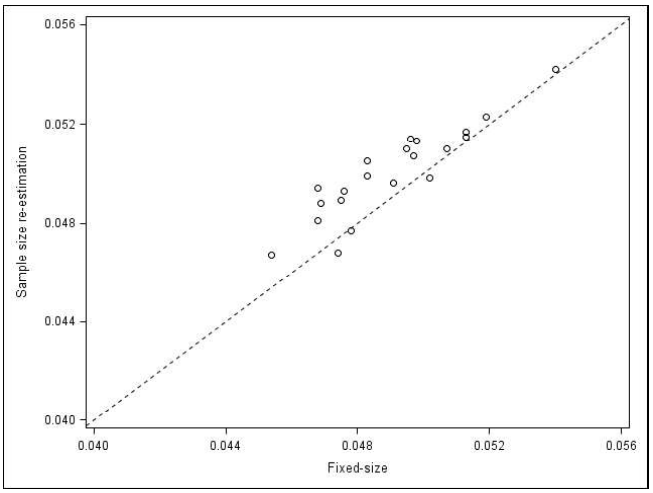


Figure 5.1-2 Type I Error (Alpha) versus ORR (With and Without BSSR) – Difference



The increase of the Type I error never excessed the above 0.052, when the Type I error is keeping the nominal level of 0.05.

Figure 5.1-3 Scatterplot of Type I Error (Alpha) in Fixed-Size and BSSR Designs - Difference



α=0.05, 1-β = 0.80, p = 0.3, 0.31 … 0.5, equivalence margin: ± 12.5%

Figure 5.1-4 Power versus ORR (With and Without BSSR) - Ratio

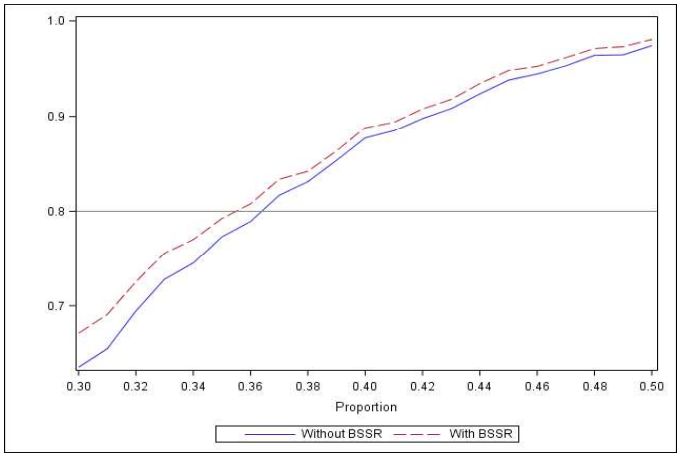
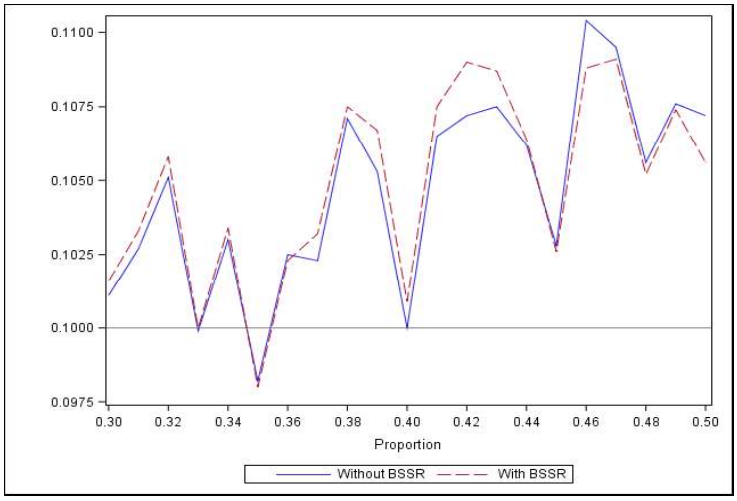
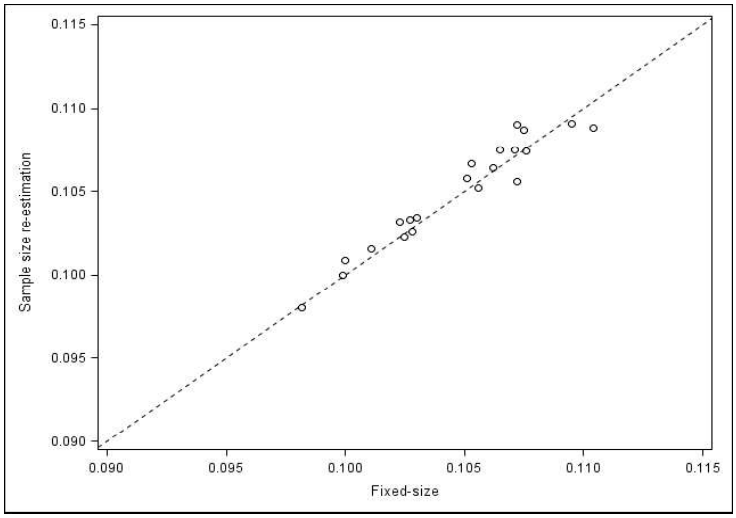


Figure 5.1-5 Type I Error (Alpha) versus ORR (With and Without BSSR) - Ratio



The simulations have been on the nominal α-level of 0.10 and simulations demonstrate a slight excess of the simulated α-level, which only slightly increased by a BSSR.

Figure 5.1-6 Scatterplot of Type I Error (Alpha) in Fixed-Size and BSSR Designs - Ratio



α=0.05, 1-β = 0.80, p = 0.3, 0.31 … 0.5, equivalence margin: (0.73, 1.36)

In conclusion, the impact on actual type I error is small in the case of a BSSR compared to the conduct without a BSSR, no adjustment of the analysis of the primary efficacy endpoint will be required.

5.2. Timing of CSR Analyses

5.2.1. Main CSR Analysis: Timing and Data Cut-off

After all randomized patients have completed Week 18 or discontinued the study, part of the sponsor clinical team will be unblinded to the randomization list, and the primary efficacy analysis will be performed. A clinical study report will be written to report the results of the primary efficacy analysis and all other efficacy and safety results up to that point in time. Data cut-off date is the date when the last randomized patient has reached or completed the tumor assessment scheduled in Week 18. Data up to Week 18 Visit will be included for the analysis of efficacy and safety endpoints. Rules for general data cut-off are included in Appendix II.

5.2.2. Final CSR Analysis: Timing and Data Cut-off

An addendum as the final study report will be written after all patients reach Week 42 or discontinued. Data cut-off date is the date when the last randomized patient has reached or completed the tumor assessment scheduled in Week 42. Data up to Week 42 Visit will be included for the analyses of secondary efficacy endpoints and safety.

5.3. Safety Monitoring

Safety will be evaluated periodically internally by Mylan Medical Monitors and Safety personnel, and externally with the Steering Committee and every 6 months with a designated data safety monitoring board (DSMB). Patient disposition, demographic, baseline, efficacy, AEs, lab assessment data will be included into for review. The details including time and frequency will be documented in DSMB Charter.

Safety blinded evaluations will be done against historical data from 5 historical published randomized trials that evaluated this combination in lung cancer and safety information from

regulatory documents of bevacizumab publicly available. The Sponsor or the Steering Committee will call an independent DSMB evaluation of unblinded data if there is any concern about the risk to the population of the study during the periodic safety evaluations, the decisions will be documented.

5.4. Patient Blinding and Unblinding

A separate PPD randomization team is assigned to generate the actual randomization schedule used for the study. An interactive web response system (IWRS) will be used to administer the randomization schedule. An unblinded pharmacist will be identified at each site, whose role will be limited to handling the study treatment. Treatment allocation via the IWRS will be provided only to the unblinded pharmacist (or authorized designee); and will be sent to the pre-specified fax/email accessible only to unblinded team members. The blinding/ unblinding team members and unblinding process will be detailed in a Blinding/ Unblinding Data Review Plan.

The unblinded pharmacist will prepare infusion bags of MYL-1402O and Avastin® and provide the infusion bags to the PI (or designee) in a blinded manner. Treatment assignment must NOT be disclosed to the PI (or designee), site or study personnel, or any sponsor representative except for the designated site monitor responsible for unblinded monitoring. The assessors (i.e., the PI and site personnel assessing safety and efficacy), the study patients as well as both the local and central radiologists must remain blinded throughout the study.

If a need for unblinding individual patient arises, one of the authorized personnel (medical monitors or site monitors) will contact IWRS and provide their User ID and PIN, the IWRS knows their role and what module they have access to. The authorized staff can then select the unblinding module and enter prompted information to unblind the patient. Any patients who are unblinded during the course of the study will be tabulated and presented in a listing.

Study unblinding will be performed at the end of period 1 when the final analysis of the primary efficacy endpoint occurs. Only individuals fulfilling select roles at Mylan and the CRO will be unblinded (see list of roles in Blinding\ Unblinding Data Review Plan). For

both parties, blinded and unblinded teams will be established prior to the completion of study period 1. Blinded teams at Mylan and the CRO will remain blinded for the duration of the study. For periods 1, 2 and the extended treatment period until the final database lock, investigators and patients will remain blinded to the treatment that the patient received.

6. General Statistical Considerations

6.1. Sample Size

To meet the different regulatory requirements, the primary endpoint will be analyzed using two approaches (ratio and difference of ORR between treatments). These different approaches lead to different sample size requirements.

FDA Regulatory Requirements

The primary efficacy analysis is based on the ratio of the MYL-1402O ORR to the Avastin® ORR at Week 18.

Following FDA's advice, four randomized clinical studies were included in a meta-analysis to estimate the treatment effect of bevacizumab in combination with chemotherapy in patients with NSCLC and to derive a similarity margin of the ratio of the ORRs. These four studies were Johnson et al, 2004; Sandler et al, 2006; Reck et al, 2010; and Niho et al, 2012. The control arm in three studies was paclitaxel plus carboplatin as background chemotherapy while in the study by Reck et al, 2010 gemcitabine were used as chemotherapy.

The p-value for testing the heterogeneity among the four selected studies was 0.230, indicating that the heterogeneity among the four studies in the meta-analysis was not significant. Hence, a fixed effects approach was used for the meta-analysis. The meta-analysis yields a risk ratio of 0.53 and an ORR for Bevacizumab plus Chemotherapy of 38%. This selection of studies and meta-analysis has been also discussed by He K at al., 2016.

Based on the meta-analysis, an equivalence margin for the risk ratio of (0.73, 1.36) was derived by maintaining 50% of the treatment effect.

Table 6.1-1 Meta-Analysis

	Control CT		Bevacizumab plus CT			
Study	(CR+PR) /N	ORR %	(CR+PR) /N	ORR %	Risk ratio	Weight (fixed)
Johnson et al	6/32	18.8	11/34	32.4	0.580	3.5%
Nishio et al	20/59	33.9	68/121	56.2	0.603	14.7%
Reck et al	71/327	21.7	114/329	34.7	0.627	37.4%
Sandler	59/392	15.1	133/381	34.9	0.431	44.4%
Meta-analysis	156/810	19.3	326/865	37.7	0.535	70%CI (0.49-0.58)

The limits for the equivalence margin have been derived using the following approach: The reverse of the upper bound of 70% CI, 1/0.58 = 1.72, indicates 72% more effect with bevacizumab. The biosimilar should retain a substantial portion of that benefit. The clinical margin representing the acceptable level of inferiority was therefore set at 50% of this 72% increase, i.e., 1.362 [= (1.72-1)/2+1]. The lower margin is the reciprocal of the upper margin, i.e., 1/1.362 =0.730.

A sample size of 588 patients (294 per treatment group) will provide 80% power for testing equivalence of MYL-1402O and Avastin® at 1-sided 5 % level of significance, for the primary endpoint ORR at Week 18. This sample size assumes that the ORR at Week 18 will be 38.0% for both MYL-1402O and Avastin®. Statistical equivalence will be declared if the 2-sided 90% confidence interval (CI) of the ratio of the 2 treatment groups falls wholly within an equivalence region defined as (0.73, 1.36).

EU/EMA Regulatory Requirements

The efficacy analysis is based on the difference of the MYL-1402O ORR and the Avastin ORR at Week 18.

The meta-analysis of the four studies results in a risk difference of -17.1% with a 70% CI of (-19.3%, -14.3%). Then equivalence margin of (-12.5%, 12.5%) maintains 65% of lower limit of the CI for the risk difference.

A sample size of 628 patients (314 per treatment group) will provide 80% power for testing equivalence of MYL-1402O and Avastin at 1-sided 2.5 % level of significance, for the primary endpoint ORR at Week 18. This sample size assumes that the ORR at Week 18 will be 38.0% for both MYL-1402O and Avastin. Statistical equivalence will be declared if the 2-sided 95% CI of the difference of the 2 treatment groups falls wholly within an equivalence margin of (-12.5%, 12.5%).

Study Sample Size Determination

The EMA requirement needs a larger sample size, therefore the planned sample size for this study will be based on size for the EMA requirement which is 628 patients (314 per treatment group). An attrition rate of 2% is assumed for the study, which leads to the number of 640 patients (320 per treatment group) to be randomized.

It is estimated that approximately 864 screened patients will yield approximately 640 patients for 1:1 randomization for having at least 628 evaluable patients, taking in to account the attrition rate of 2%. As described in section 5.1, a blinded interim analysis with aim of a sample size re-estimation will be performed after at least 30% of the patients have completed week 18. This blinded sample size re-estimation (BSSR) may have an effect on the actual on type I error (Friede T, Kieser M, 2004). The magnitude of the Type I error increase was investigated by performing simulations for the situation with and without BSSR. Results from the simulations suggest that the impact of a BSSR on the actual type I error is small. Therefore, a Type I error adjustment has not been considered for the sample size estimation.

6.2. Randomization and Stratification

Patients who meet all of the inclusion and none of exclusion criteria and who are approved by the medical monitor will be considered eligible for entry into the study and will be randomly assigned to treatment in the study within 3 days prior to Day 1 of Cycle 1.

Approximately 640 eligible patients (320 patients per group) from approximately 190 sites worldwide will be randomly assigned to one of 2 treatment arms (MYL-1402O or

Avastin®). An IXRS will be used to administer the randomization schedule. The randomization schedule is generated using SAS software Version 9.2 or later (SAS Institute Inc, Cary, North Carolina) for IXRS, which will link sequential patient randomization numbers to treatment kit numbers. The randomization schedule will be stratified by gender (male or female), smoking status (smoker or <100 cigarettes in entire lifetime), and number of metastasis sites (one site or multiple sites). A listing of randomization scheme and codes including patient identification, stratum, and treatment assigned will be created.

6.3. Reporting Conventions

Unless otherwise stated, all table summaries will be presented by treatment arm and overall.

For continuous variables, descriptive statistics will typically include the number of patients, mean, standard deviation (SD), median, minimum, and maximum. For categorical variables, descriptive statistics will typically include the number and percentage of patients in each category. Unless otherwise noted, data from all sites will be combined in the computation of these descriptive statistics.

The percentage will be suppressed when the count is zero in order to draw attention to the non-zero counts. The denominator for all percentages will be the number of patients within the analysis set of interest, unless otherwise specified.

In statistical summaries, mean and median will be displayed to one more decimal place than the original value; minimum and maximum will keep the same number of decimal places as the original value; SD will be displayed to two more decimal places than the original value; percentages will generally be reported to one decimal place.

The statistical test for the primary efficacy endpoint will be performed using a 5% significance level (1-sided), leading to a 90% (2-sided) CI for FDA requirement, while using a 2.5% significance level (1-sided), leading to a 95% (2-sided) CI for EMA requirement.

For the secondary efficacy endpoints, no formal hypothesis testing will be performed. The time to event secondary endpoints will be analyzed using log-rank tests.

Patients will be identified in the listings by patient identification number concatenated with the site number.

All statistical analyses will be performed using SAS® version 9.2 or higher.

6.4. General Methods and Definitions

Conversion of Days to Weeks, Months, and Years

The following conversion factors will be used to convert days into weeks, months, or years: 1 week = 7 days, 1 month = 30.4375 days, and 1 year = 365.25 days.

Data after Cut-off Date

Data after cut-off date will not be used for analysis purposes or included in any listings.

Definition of Baseline

Baseline is defined as the last non-missing measurement taken prior to the first dose of interventional therapy (either MYL-1402O, Avastin®, carboplatin, or paclitaxel) for safety analyses and as screening visit for tumor assessment per RECIST v1.1.

Definition of Duration

Duration will be calculated as the difference of start and stop date + 1, if not otherwise specified.

Definition of Study Day

Study day = event/assessment date – date of first administration of IMP +1 for assessment performed on or after the first dose date.

Study day = event/assessment date – date of first administration of IMP for assessment performed prior to the first dose date.

Study Day 1 corresponds to the date of first administration of IMP and Study Day -1 corresponds to the day before the first administration of IMP. If patients were randomized but not dosed, randomization date will be used instead.

Definition of Treatment Period

The treatment period is defined as the time from the first administration of IMP to the last administration of IMP. This period will include 18 weeks in Period 1 (bevacizumab in combination with carboplatin and paclitaxel), 24 weeks in Period 2 (bevacizumab monotherapy), and the Extended Treatment Period (continued bevacizumab monotherapy until PD, death, AE, withdrawal of consent, or discontinuation from study drug).

Handling of Missing Data

All data will be evaluated as reported and no imputation of missing values will be done unless otherwise stated. Missing statistics (i.e. when they cannot be calculated) will be presented as a single dash ("NC"). For example, if n=1, the measure of variability SD cannot be computed and will be presented as "NC".

Unscheduled Visits

Data collected at unscheduled visits will not be included for by-visit summaries of data but will contribute to analyses in which post-baseline data is considered (e.g. worst post-baseline value) and analyses of efficacy endpoints.

6.5. Analysis Sets

The following analysis sets are defined for the statistical analyses. If a patient was randomized more than once, the patient will be included for the analysis based on the randomized treatment the patient actually entered into the study.

6.5.1. Screened Set

The Screened set will consist of all patients who have completed an informed consent form.

6.5.2. Intent-to-Treat (ITT) Set

The ITT set will consist of all randomized patients. All analyses using the ITT set will group patients according to randomized treatment. The ITT set will be used for efficacy and other analyses including immunogenicity.

6.5.3. Modified Intent-to-Treat Set (Only Used for BSSR Interim Analysis)

The mITT set will consist of all randomized patients who complete at least one dose of MYL-1402O or Avastin®, patients who met one of the following conditions will be excluded from mITT set:

- No assessable post-baseline tumor scan per independent review, no treatment-emergent AEs, with <= 2 doses of MYL-1402O or Avastin®, and discontinued from study due to reasons other than death; or
- Randomized and dosed, but later found to be ineligible due to inclusion criterion 6 or exclusion criteria 7, 9, 10, 18, and reported as protocol deviation in CTMS.

The mITT set will be only used for BSSR interim analysis.

6.5.4. Per Protocol (PP) Set

The PP set will consist of all randomized patients who complete at least one dose of MYL-1402O or Avastin® and do not have protocol deviations with significant impact on the (study) endpoints during the study. Major protocol deviations will be defined in Section 7.2. Exclusion of patients from the PP set due to major protocol violations will be decided in a Blinded Data Review (BDR) meeting prior to study unblinding. All analyses using the PP set will group patients according to randomized treatment. The PP set will be used for the sensitivity analyses of the efficacy endpoints.

6.5.5. Safety Set (SS)

The SS will consist of all randomized patients who complete at least one dose or partial dose of MYL-1402O or Avastin®. All analyses using the SS will group patients according to treatment actually received. The SS will be used for the safety analyses and the immunogenicity analysis.

6.5.6. Population PK (PopPK) Set

The PopPK set will consist of all randomized patients who complete at least 1 dose of allocated study medication and who provide at least 1 evaluable post-dose drug

concentration for PopPK analysis. All analyses using the PopPK set will group patients according to IMP actually received.

6.5.7. Subgroups to be Analyzed

The analysis of primary and secondary endpoints will be done for the following subgroups, if applicable:

- Age (<65 years or ≥65 years)
- Gender (Male or Female) per CRF
- Race (White or non-White)
- Smoking status (smoker, defined as >=100 cigarettes in entire lifetime, or non-smoker, defined as <100 cigarettes in entire lifetime) per CRF
- Number of metastasis sites (1 or multiple) per IWRS: one metastasis site if disease stage is M1A or M1B; multiple metastasis sites if disease stage is M1C at screening.
- Prior radiation therapy (yes/no)
- Baseline ECOG performance status (0 or 1)
- EGFR mutation status (unknown or negative)
- EML4-ALK alterations (unknown or negative)
- Region: Europe, India, Southeast Asia

6.6. Treatment Misallocations

Patients may receive incorrect IMP during the study due to unforeseeable reasons, those patients will be considered as major protocol deviation and thus excluded from PP set for the analyses.

6.7. Handling of Incomplete Dates

Birth Date: Partial birth date will be imputed as follows for the analyses:

- If the birth date month is not missing and the birth date day is missing, the birth date will be imputed as the first day of the month (01MONYYYY).
- If the birth date day and month are both missing, the birth date will be imputed as the first day of the year (01JanYYYY).

AE Onset/End Date: For the purpose to identify treatment emergent adverse event (TEAE), partial AE onset and end dates will be imputed as follows:

Partial onset dates (where UK and UKN indicate unknown or missing day and month respectively):

- UK-MMM-YYYY: If the month and year are different from the month and year of the first dose of IMP, assume 01-MMM-YYYY. If the month and year are the same as the first dose of IMP month and year, and the end date (after any imputation) is on or after the first dose of IMP, then assume the date of the first dose of IMP. If the month and year are the same as the first dose of IMP month and year, and the end date (after any imputation) is prior to the first dose of IMP, then assume the end date for the onset date.
- DD-UKN-YYYY/UK-UKN-YYYY: If the year is different from the year of first dose of study drug, assume 01-JAN-YYYY of the collected year. If the year is the same as the first dose of IMP year, and the end date (after any imputation) is on or after the first dose of IMP, then assume the date of the first dose of IMP. If the year is the same as the first dose of IMP, and the end date (after any imputation) is prior to the first dose of IMP, then assume the end date for the onset date.

Partial end dates:

- UK-MMM-YYYY: Assume the last day of the month if not resulting in a date later than the date of the patient's death; in the latter case, the date of death will be assumed.
- DD-UKN-YYYY/UK-UKN-YYYY: Assume 31-DEC-YYYY if not resulting in a date later than the date of the patient's death; in the latter case, the date of death will be assumed.

When the date is completely missing, no imputation will be performed and the AE will be considered as treatment emergent, unless there is rational to clarify otherwise, eg. AE end date is prior to the first dose date.

Start/End Dates of Prior/Concomitant Medication: For the purpose of inclusion in prior and/or concomitant medication tables, partial medication start and stop dates will be imputed as follows:

Partial start dates:

- UK-MMM-YYYY: If the month and year are different from the month and year of the first dose of IMP, assume 01-MMM-YYYY. If the month and year are the same as the first dose of IMP month and year and the end date (after any imputation) is on or after the first dose of IMP, then assume the date of the first dose of IMP. If the month and year are the same as the first dose of IMP month and year and the end date (after any imputation) is prior to the first dose of IMP, then assume the end date for the start date;
- DD-UKN-YYYY/UK-UKN-YYYY: If the year is different from the year of the first dose of IMP, assume 01-JAN-YYYY of the collected year. If the year is the same as the first dose of IMP year and the end date (after any imputation) is on or after the first dose of IMP, then assume the date of the first dose of IMP. If the year is the same as the first dose of IMP and the end date (after any imputation) is prior to the first dose of IMP, then assume the end date for the start date.

Partial stop dates:

- UK-MMM-YYYY: Assume the last day of the month if not resulting in a date later than the date of the patient's death; in the latter case, the date of death will be assumed;
- DD-UKN-YYYY/UK-UKN-YYYY: Assume 31-DEC-YYYY if not resulting in a date later than the date of the patient's death; in the latter case, the date of death will be assumed.

If the start or stop date is completely missing, no imputation will be performed and the determination of prior medication or post-medication will be based on non-missing stop or start date, respectively; otherwise, the medication will be considered as concomitant.

Dates for Disease History: Missing or partial dates for disease history (initial diagnosis date, first diagnosis of metastatic or advanced disease) will be imputed as following:

- If the day is missing, it will be imputed to the 15th day of the month; if both day and month are missing and the year is prior to the year of the first IMP (randomization date if not dosed), the month and day will be imputed as July 1st; if both day and month are missing and the year is same as the year of the first IMP (randomization date if not dosed), the month and day will be imputed as January 1st. If the date is completely missing, no imputation will be performed. If imputed first diagnosis date of metastatic or advanced disease date is prior to or on initial diagnosis date, set it to the date of initial diagnosis date if initial disease stage is stage III or IV, or set it to the date of initial diagnosis date +1 if initial disease stage is not stage III or IV

Dates for Prior Anti-Cancer Drug Therapy: Partial dates for prior anti-cancer drug therapies will be imputed as following:

- If the day is missing, it will be imputed to the first day of the month; if both day and month are missing, no imputation will be performed.

6.8. Multiplicity Adjustments

No formal adjustment for multiplicity will be performed.

7. Patient Disposition

7.1. Disposition

Patient enrolment by country and center will be summarized for each treatment arm and overall for screened patients. A summary of the analysis sets includes the number of patients for the following categories: patients in the ITT set, patients in the Safety set, and patients in the PP set by treatment arm and overall.

Patient disposition will be summarized for all randomized patients. A disposition of patients includes the number and percentage of patients who were baseline failures, patients who discontinued IMP , and reasons for IMP discontinuation by treatment arm and overall. Baseline failure is defined as any patient that meets initial eligibility requirements, is randomized to treatment, however does not receive first dose of bevacizumab and is deemed ineligible for continuation in the study. The reason for IMP discontinuation may include Adverse Event, Disease Progression, Death, Lose to Follow-Up, Withdrawal Consent, PI Decision, Protocol Violation, and Administrative Reason by Sponsor etc.

Study completer is defined as follows:

- If patient terminates treatment prematurely for any reason but still continues participation in Survival Follow-Ups in the study until week 42, the patient is a study completer.
- If a patient continues treatment beyond week 42, the patient is a study completer.

If a patient terminates treatment prematurely for any reason and stops participation in the study before week 42 for any reason (withdrawn consent, lost to follow-up, etc.), the patient is a study early terminator. Study completers will be derived from clinical database such as Drug Administration, Per Time Point Visit Response, and Survival eCRF pages while study early terminators will be extracted from Study Completion eCRF page. The reasons for study early termination will also be summarized in this table and categorized as one of the following: Baseline Failure or Not Treated, Death, Lost to Follow-up, Withdrawal Consent, Administrative, and Other. The percentages will be based on the number of patients randomized.

Patient disposition data will be presented in a listing including treatment arm, patient ID, first/last dose of IMP/CP, baseline failure (yes/no), date/reason for EOT, date/reason for study early termination.

7.2. Protocol Deviations

A major and/or serious protocol deviation (otherwise known as a protocol violation) is one that materially affects the safety of the patients and/or the evaluation of primary or key secondary efficacy endpoints of the study.

Current ICH and EU GCP guidelines list the major protocol deviations that must be listed in the clinical report. These include:

- patients that are dosed on the study despite not satisfying the inclusion criteria;
- patients that develop withdrawal criteria whilst on the study but are not withdrawn;
- patients that receive the wrong treatment or an incorrect dose;
- patients that receive a prohibited concomitant medication.

Specifically for this study, major protocol deviation will be assessed based on the following criteria.

[1]	No Written Informed Consent provided	Inclusion Criteria	Major
[2]	Was not randomized in the study but dosed	Treatment Schedule	Major
[3]	Did not receive correct treatment as randomized	Treatment Schedule	Major
[4]	No measurable disease confirmed by independent review from Parexel.	Inclusion Criteria	Major
[5]	Did not complete 18 weeks of study and at least 4 doses of chemotherapy plus MYL 1402O or Avastin and at least one tumor assessment with the exception of progression or death or AE	Assessment Schedule	Major
[6]	Forbidden prior therapy/medication	Inclusion Criteria	Major/Minor
[7]	Violation of other inclusion criteria	Inclusion Criteria	Major/Minor
[8]	Violation of other exclusion criteria	Exclusion Criteria	Major/Minor
[9]	Use of prohibited concomitant medication.	Concomitant Medication	Major/Minor
[10]	Previous or current medical conditions.	Exclusion Criteria	Major/Minor

Major protocol deviations will be identified based upon the eCRF database and determined for patients by medical review or clinical monitoring. The assignment of major or minor for some criteria mentioned above will be decided on a case-by-case basis. Major protocol deviations and patients to be excluded from PP set due to major protocol deviations will be

finalized during a BDR meeting prior to data un-blinding for the analysis. A summary table of patients to be excluded from PP set and SS will be provided. Patients to be excluded from PP set and SS will be presented in a data listing.

Protocol deviations will also be classified as critical, major, minor from clinical perspective, critical and major protocol deviations will be summarized in a table, all subject level protocol deviations will be presented in a data listing. Protocol deviation stated in this SAP contains ICH-GCP deviations.

8. Demographics and Baseline Characteristics

8.1. Demographics

The demographics and baseline characteristics will be summarized using descriptive statistics for the following variables for ITT and PP sets:

- Age (in years)
- Age category (<65/≥65 years)
- Gender (Male, Female)
- Race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, and Other)
- Race categories (White vs. non-White)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Height (cm)
- Weight (kg)
- Body Surface Area (BSA) (m^2)
- Body Mass Index (BMI) (kg/m^2)
- ECOG performance status

- Smoking status (Smoker or Non-Smoker) per CRF
- Number of metastasis sites (one or multiple) per IWRS
- Region (Europe, India, Southeast Asia)

Baseline ECOG, weight and height will be the last non-missing values prior to the first dose of interventional therapy or randomization date if not dosed. Age (years), baseline height (cm), baseline weight (kg), BSA (m^2), and baseline BMI (kg/m^2) will be summarized using descriptive statistics such as mean, SD, median etc. The number and percentage of patients by age category, sex, race, race category, ethnicity, ECOG performance status, and smoking status will be reported using frequency count and percentage. Percentages will be based on the total number of patients in the applicable analysis set.

Age, BSA, and BMI will be derived as:

- Age (year) = (date of informed consent – date of birth + 1)/365.25.
- BSA (m^2) = SQRT(weight (kg) × height (cm))/60
- BMI (kg/m^2) = weight(kg)/[height(m)]^2.

The integer part of the calculated age will be used for reporting purpose.

The listing of demographics and baseline characteristics will include the following information: treatment arm, patient identifier, age, sex, race, ethnicity, baseline height (cm), weight (kg), BSA (m^2), and BMI (kg/m^2), smoking status, number of metastasis sites, and baseline ECOG.

8.2. Baseline Disease Characteristics

Disease histories are collected on Initial Disease Diagnosis electronic Case Report Form (eCRF) page. Partial date will be imputed as described in Section 6.7. The disease history table will include descriptive statistics for the following variables by treatment arm and overall using ITT set:

- Tumor histology

  - Adenocarcinoma
  - Large cell Carcinoma
  - Bronchoalveolar
  - Not other specified
  - Other
- Disease stage at initial diagnosis
  - Stage I
  - Stage II
  - Stage III
  - Stage IV
  - Unknown
- Disease sub-stage for stage IV patient
  - M1A
  - M1B
  - M1C
- Time since initial diagnosis of disease (months), defined as (the randomization date – the date of first diagnosis)/30.4375
- Time since initial diagnosis of metastatic or advanced disease (months), defined as (the randomization date – the date of first diagnosis of metastatic or advanced disease)/30.4375
- EGFR mutation (negative, unknown, missing)
- EML4-ALK alternation (negative, unknown, missing)

Listing of disease history will be provided with all relevant data (tumor histology, initial diagnosis date and disease stage, first diagnosis of metastatic or advanced disease, disease sub-stage, mutations) and derived variables used in the above table.

8.3. Stratification Factors

The stratification factors for randomization contain gender (male or female), smoking status (smoker (>=100 cigarettes in entire lifetime) or non-smoker (<100 cigarettes in entire lifetime)), and number of metastasis sites (one site or multiple sites). All stratified factors will be included for subgroup analyses of efficacy endpoints.

If a subject is randomized under an incorrect stratum, i.e. mis-stratified by gender, smoking status and/or number of metastasis sites, the respective observed stratums will be used in all analyses. A list of subjects that are mis-stratified will be provided including gender from CRF and IWRS, smoking status from CRF and IWRS, # of metastatic sites from IWRS. Gender and smoking status per CRF and # of metastatic sites would be used consistently for all outputs and analyses except randomization listing which will display strata per IWRS.

8.4. Medical History

Medical history will be coded using the most recent version of Medical Dictionary for Regulatory Activities (MedDRA). The number and percentage of patients with any medical history will be summarized for each system organ class (SOC) and preferred term (PT) by treatment arm and overall using ITT set. Percentages will be calculated based on number of patients in the applicable analysis set.

Each patient will be counted only once within each PT or SOC. Medical history will be sorted in descending order of frequency (total column) by SOC and by PT within each SOC, unless otherwise stated. The dictionary version used for reporting the study will be described in the relevant table and listing footnotes.

Listing of medical history data by patient will include coded terms and all the relevant data fields as collected on the Medical History eCRF page.

8.5. Physical Examination

Patients with abnormal physical examination at screening will be tabulated by examination criteria and by treatment arm and overall using ITT set.

Listing of physical examination at screening will include patient ID, examination criteria, examination result.

8.6. Inclusion and Exclusion Criteria

Inclusion and exclusion criteria will be presented in a data listing for all screened patients and summarized by failing criteria for all screening failures.

9. Treatments and Medications

9.1. Prior and Concomitant Medications/Procedures

Use of prior or concomitant medications will be recorded in the patient's eCRF. The minimum requirement is that drug name and the dates of administration are to be recorded. This will include all prescription drugs, herbal products, vitamins, minerals, and over-the-counter medications. Any changes in concomitant medications will also be recorded in the patient's eCRF. Prior and concomitant medications will be coded according to the latest version of World Health Organization drug dictionary (WHO DRUG). Partial or missing start and stop dates will be imputed per Section 6.7 for the purpose of analysis. Original dates will be displayed in the corresponding listing. A patient will be counted only once within a given drug class and within a given drug name for the summary, even if he/she received the same medication at different times. The summary will be sorted on decreasing frequency (total column) of drug class and drug name in a given drug class. In case of equal frequency, alphabetical order will be used.

9.1.1. Prior Medications

A prior medication is defined as any medication with a stop date prior to the first dose of IMP (randomization date if not dosed with IMP). The total number of prior medications and the number and percentage of patients with at least one prior medication will be summarized

by treatment arm and overall. The number and percentage of all prior medications will be summarized by Anatomical Therapeutic Chemical (ATC) level 2 and PT and by treatment arm and overall based on ITT set. Percentages will be calculated based on number of patients in the ITT set.

9.1.2. Concomitant Medications

Concomitant medications will be any drug taken at any time from Day 1 of Cycle 1, through the date of last dose of interventional therapy (including Period 1, Period 2, and Extended Treatment Period), other than the interventional therapy (including bevacizumab, carboplatin, paclitaxel) and premedication. For patients randomized but not dosed with interventional therapy, medications used on or after randomizations date are considered as concomitant medications.

The total number of concomitant medications and the number and percentages of patients with at least one concomitant medication will be summarized by treatment arm and overall. The number and percentages of all concomitant medications will be summarized by treatment arm and overall and listed by ATC level 2 and PT. Summaries will be performed using the ITT set. Percentages will be calculated based on number of patients in the ITT set.

Any medication (except subsequent anti-cancer treatments) started after last dose of interventional therapy will be considered as post-medication. All the medications will be displayed in a data listing with a flag to identify prior, concomitant, or post medication. Subsequent anti-cancer therapy will be summarized in Section 9.1.4.

9.1.3. Prior Anti-cancer Treatments

The prior anti-cancer treatments and procedures are collected under the Prior Anti-Cancer Drug Therapies, Prior Anti-Cancer Radiotherapies and Prior Anti-Cancer Surgeries eCRF pages.

The overall summary of prior anti-cancer treatments will include: the number and percentage of patients by type of prior anti-cancer treatment, i.e.

- Number of patients with at least one type of prior anti-cancer treatment
- Number of patients with at least one prior anti-cancer drug therapy
- Number of patients with at least one prior anti-cancer radiotherapy
- Number of patients with at least one prior anti-cancer surgery

Summary of prior anti-cancer drug therapy will include the following variables by treatment arm and overall:

- Number of patients with at least one prior anti-cancer drug therapy
- Intent of therapy: adjuvant / neo-adjuvant
- Primary reason therapy was discontinued (disease progression/ maximum clinical benefit/ toxicity/ completed treatment/ other)

The listings of prior anti-cancer treatments and procedures will also be provided: a) listing of prior anti-cancer drug therapy, b) listing of prior anti-cancer radiotherapy, and c) listing of prior anti-cancer surgery. These will include patient identifier and all the relevant collected data-fields on the corresponding eCRF pages such as therapy start/stop date, intent of therapy, reason therapy was stopped, PD date and documentation of PD for anti-cancer drug therapy.

9.1.4. On-Study and Subsequent Anti-Cancer Treatment

On-study and subsequent radiotherapy will be presented in a data listing including variables such as body site, cumulative dose and unit, therapy start/stop date and intent. On-study and subsequent surgery will be presented in a data listing including variables such as date and type of surgery, and tumor involved. Anti-cancer drug therapy after discontinuation of interventional therapy will be provided in a data listing with all relevant information such as patient identifier, drug name, start date, end data, dose, dose unit etc.

9.2. Study Treatments

IMP (MYL-1402O or Avastin®) will be administered by i.v. infusion at a dose of 15 mg/kg every 21days (±3days) for up to 6 treatment cycles (1 treatment cycle = 21 days based on bevacizumab administration) in combination with CP and then alone as a maintenance. The IMP exposure and compliance will be summarized for the Safety set by treatment arm. Administration of IMP and CP will be presented in the listing including name of interventional therapy, start/ stop date/time, study day, relative study day to previous administration, most recent body weight prior to infusion, actual and planned dose (kg), dose delay/missed (yes or no), reason for missing dose etc.

9.2.1. Extent of IMP/CP Exposure

Analysis of IMP/CP exposure will be based on the actual dose administered (mg) or actual dose level (actual dose administered/ weight, mg/kg). The last available weight of the patient on or prior to the day of dosing will be used for the calculation.

The summary of IMP/CP exposure based on the Safety set will include the following variables per patient (a cycle refers to the planned dosing interval of three weeks):

- Treatment duration (in weeks), defined as (the last dose date – the first dose date + 21)/7
- Number of administrations.
- Cumulative dose (mg/kg), defined as sum of actual dose level for IMP, or cumulative dose (mg), defined as sum of actual dose for CP.
- Dose intensity (mg/kg/cycle), defined as sum of actual dose levels (mg/kg) / (treatment duration (week)/3) for IMP, or dose intensity (mg/cycle), defined as sum of actual dose (mg) / (treatment duration (week)/3) for CP.
- Relative dose intensity (%), defined as dose intensity (mg/kg/cycle) * 100/ planned dose level (mg/kg/cycle) for IMP, or defined as cumulative dose (mg) * 100/sum of planned dose (mg) for CP, the planned dose level is 15 mg/kg for IMP based on the study protocol.
- Relative dose intensity by the following categories:

- >0.9
- >0.8-0.9
- <=0.8

A listing of treatment exposure and compliance will include treatment arm, patient identifier, and above derived variables summarized in the table.

9.2.2. IMP Compliance and Modifications

Per protocol, patients will receive the IMP once every 3 weeks +/- 3 days. Dose delays will be grouped into the following categories based on the deviation of the actual to the planned treatment administration day (relative to the previous non-zero dose date): no delay (including 1-3 days delays), 4-7 days delay, 8 or more days delay. For example, if one patient receives the study drug on day 1, then the next study drug administration date will be on day 22; however, if the patient receives the study drug at day 23, 24 or 25, this is considered as 'no delay'. Any zero dose prior to the last treatment administration is considered as a dose interruption.

The summary of dose delays will be based on the Safety set and includes the following categories:

- No delay
- 4-7 days delay
- 8 or more days delay

The categorization is based on the maximum length of delay, i.e. the worst case of delay if patients have multiple dose delays.

A listing containing patient identifier, visit, and unique study drug batch ID will also be created.

10. Efficacy Analysis

Efficacy will be assessed based on tumor response and survival. Tumor assessments are to be conducted every 6 weeks (±3 days) up to 18 weeks, independent of delays in treatment administration in Period 1 and every 12 weeks (±3 days) independent of delays in treatment administration after 18 weeks and beyond Period 2 and Extended Treatment Period. Tumor response will be assessed using RECIST1.1 criteria locally by investigators and centrally by independent reviewers (radiologists). Time point overall response and/or best overall response (BOR) up to week 18 or 42 and associated dates will be assessed by investigator and independent reviewers (radiologists). BOR derived from time point overall responses per independent reviewers up to Week 18 will be used for primary analysis of primary efficacy endpoint, the derivation is to choose best response in the order of CR, PR, Stable Disease, PD and NE among time point overall responses up to Week 18 for Week 18 analysis or Week 42 for Week 42 analysis per RECIST 1.1. Secondary efficacy endpoints such as DCR, DOR, and time to response at Week 18 and 42 will also be based on derived BORs, confirmation of CR or PR will be conducted and incorporated into BOR by independent reviewers, which will be used as a sensitivity analysis of primary efficacy endpoint at Week 18. To summarize, four BORs will be used for efficacy analyses:

- BOR derived from independent reviewer data among time point overall responses for primary analyses per RECIST 1.1.
- BOR from independent reviewers with confirmation of CR/PR used for sensitivity analyses.
- BOR from investigator used for sensitivity analyses
- BOR derived from time point overall response from investigator data used for sensitivity analysis

Number and percentage of patients achieving BOR = CR, PR, Non-CR/Non-PD, stable disease, PD will also be summarized under corresponding categories. If patients had baseline and at least one post-baseline lesion assessment, and their BOR was assessed as NE, these

patients will be summarized under NE; if patients had baseline lesion assessment and no post-baseline lesion assessment, these subjects will be categorized as not done (ND) and summarized as such.

10.1. Primary Efficacy Endpoint

The primary efficacy endpoint is the ORR based on best tumor response as assessed by an independent reviewer at any time point during the first 18 weeks, according to RECIST1.1. The ORR is defined as the proportion of patients with CR or PR as the BOR during the first 18 weeks, the number of patients in the ITT set will be used as the denominator for the calculation. Patients with response that cannot be evaluated per RECIST1.1 or without any tumor assessment prior to or on Week 18 will be scored as non-responders. The primary efficacy endpoint is analyzed differently to meet the requirements of the FDA and the EMA. It is expected that both approaches will lead to congruent conclusions, although the results will be discussed separately for the two requirements to a certain extent. Therefore no adjustment of the type I error is done for the primary analysis.

10.1.1. Efficacy Endpoint Analysis to Meet the FDA Requirements

The ORRs based on derived BOR up to week 18 for both MYL-1402O and Avastin® will be calculated, the ratio of the ORRs will be used to determine if MYL-1402O is equivalent to Avastin®. A 2-sided 90% CI for the ratio of the ORRs at week 18 will be calculated based on the method of logarithmic transformation (OPTION=RELRISK in TABLES statement in PROC FREQ) with no adjustment for covariates (Blackwelder 1993). This primary endpoint analysis will be conducted using the ITT set.

The statistical hypotheses associated with the primary analysis of ORR at week18 are as follows:

- H_0: ($ORR_{MYL\text{-}1402O}$ / $ORR_{Avastin} \leq 0.73$) or ($ORR_{MYL\text{-}1402O}$/ $ORR_{Avastin} \geq 1.36$)
- H_1: $0.73 < ORR_{MYL\text{-}1402O} / ORR_{Avastin} < 1.36$,

where $ORR_{MYL\text{-}1402O}$ and $ORR_{Avastin}$ are the ORRs for MYL-1402O and Avastin®, respectively.

The use of a 90% CI to evaluate equivalence is the same as performing two 1-sided tests at an alpha level of 0.05 for each test.

10.1.2. Efficacy Endpoint Analysis to Meet the EMA Requirements

The ORRs calculated in Section 10.1.1 will be analyzed to meet the EMA requirement, an asymptotic 2-sided 95% CI for the difference in ORRs at Week 18 will be calculated using Wald confidence interval (OPTION=RISKDIFF in TABLES statement in PROC FREQ) without adjustment for covariates.

The statistical hypotheses associated with the primary analysis of ORR at Week 18 for EMA are as follows:

- H_0: ($ORR_{MYL\text{-}1402O}$-$ORR_{Avastin}$≤ -12.5%) or ($ORR_{MYL\text{-}1402O}$-$ORR_{Avastin}$≥ 12.5%)
- H_1: -12.5%< ($ORR_{MYL\text{-}1402O}$-$ORR_{Avastin}$) < 12.5%,

Where $ORR_{MYL\text{-}1402O}$ and $ORR_{Avastin®}$ are the ORR for MYL-1402O and Avastin®, respectively.

10.1.3. Sensitivity Analyses

Sensitivity analysis of the primary endpoint will be performed based on the following:

- PP set

The analysis will be conducted using the same method for primary endpoint specified in Section 10.1.1 or 10.1.2 based on PP set using derived BOR from independent reviewer. The number of patients in the PP set will be used as the denominator for the calculation of ORR.

- Investigator's assessment of tumor response, derived BOR from investigator data, and BOR confirmed by independent reviewer

The analysis will be conducted using the same method for primary endpoint specified in Section 10.1.1 or 10.1.2 but based on tumor response data from investigator's assessment, derived BOR from investigator data, and confirmed BOR from independent reviewer for ITT set.

- Cochran-Mantel-Haenszel (CMH) test adjusted for stratification factors (OPTION=CMH relrisk commonriskdiff in TABLES statement in PROC FREQ), common relative risk and risk difference and associated 95% CI will be reported. This analysis will be based on derived BOR from independent reviewer.

10.2. Secondary Efficacy Endpoints

10.2.1. Disease Control Rate

DCR is defined as number of patients with disease under control (BOR of CR, PR, or stable disease) divided by the number of patients in the analysis set. Patients with BOR of non-CR/non-PD are not considered as having achieved objective response, not clinical benefit, i.e. they will only be counted in the denominator of the rate, but not the numerator. The ratio of the DCRs and its 90%CI will be estimated using the method of logarithmic transformation with no adjustment for covariates, same method used in Section 10.1.1. An asymptotic 2-sided 95% CI for the difference in DCRs will be also calculated, same method used in Section 10.1.2. These analyses will be based on ITT set using BOR derived from independent reviewer data. Sensitivity analyses will be performed based on PP set and investigator data similar to those analyses for ORR as described in Section 10.1.3. Common relative risk, risk different and associated 95% CIs will be reported based CMH test adjusted for stratification factors.

10.2.2. Progression-Free Survival

PFS time is defined as the time from randomization to the date of the first documentation of PD or death by any cause, whichever occurs first. PFS will be censored in the following scenarios in Table 10.2.2-1:

Table 10.2.2-1. Censoring Scheme for Primary PFS Analysis

Situation	**Date of Progression/Censoring**	**Outcome**
Incomplete or no baseline tumor assessment	Date of randomization	Censored

No adequate post-baseline tumor assessment and no death	Date of randomization	Censored
No progression and no death	Date of last tumor assessment with no documented progression	Censored
New anticancer treatment started without a prior reported progression	Date of last tumor assessment prior to or on the date of start of new treatment	Censored
Progression documented between scheduled visits	Earliest of: Date of tumor assessment showing new lesion; Date of last tumor assessment (target or non-target) with documented progression	Progressed
Death without progression and without subsequent anti-cancer therapy	Date of death	Progressed

PFS = (date of PD or death/censoring - date of the randomization + 1)/7 (weeks).

Median and 95% CI will be reported in months, this applies to all the time to event analyses such as PFS, DOR and OS. The Kaplan-Meier method will be used to estimate parameters for PFS based on ITT set. Kaplan-Meier estimates (product-limit estimates) will be presented together with a summary of associated statistics including corresponding two-sided 95% CIs, p-values will be based on the log-rank test. In particular, survival rates at weeks 18, 30, and 42 will be estimated with corresponding two-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley (Brookmeyer R, et al, 1982) and CIs for the survival function estimates at the above defined time points will be derived using the log-log transformation according to Kalbfleisch and Prentice (Kalbfleisch JD et al., 1980) (CONFTYPE = loglog default option in SAS PROC LIFETEST). The estimate of the standard error will be computed using Greenwood's formula. Kaplan-Meier plots by treatment arm will be produced for both PFS. In addition, a Cox proportional regression model will be used to estimate the hazard ratio (HR) and the corresponding 95% CI for the treatment effect. The main Cox proportional regression model will have treatment

and no covariates. Sensitivity analyses will be performed with covariates added to the main model, the covariates to be included are stratification factors, race, age, prior radiotherapy, and baseline ECOG performance status.

Sensitivity analysis for PFS using the Kaplan-Meier method will performed based on the following:

- PP population
- Investigator's assessment of tumor response
- PFS accounting for subsequent therapy (censoring scheme in Table 10.2.2-2). Patients will be defined similarly to the primary definition except that subsequent anti-cancer therapy will be considered as an event.

Table 10.2.2-2. Sensitivity Analysis of PFS Accounting for Subsequent Therapy

Situation	**Date of Progression/Censoring**	**Outcome**
Incomplete or no baseline tumor assessment	Date of randomization	Censored
No adequate post-baseline tumor assessment and no death	Date of randomization	Censored
No progression and no death	Date of last tumor assessment with no documented progression	Censored
New anticancer treatment started without a prior reported progression	Date of last tumor assessment prior to or on the date of start of new treatment	Progressed
Progression documented between scheduled visits	Earliest of: Date of tumor assessment showing new lesion; Date of last tumor assessment (target or non-target) with documented progression.	Progressed
Death without progression and without subsequent anti-cancer therapy	Date of death	Progressed

- PFS accounting for missing tumor assessment prior to PFS event (censoring scheme in Table 10.2.2-3). Patients will be defined similarly to the primary definition except that death or PD after more than one missed tumor assessment will not be considered as an event.

Table 10.2.2-3. Sensitivity Analysis of PFS Accounting for Missing Tumor Assessment

Situation	**Date of Progression/Censoring**	**Outcome**
Incomplete or no baseline tumor assessment	Date of randomization	Censored
No adequate post-baseline tumor assessment and no death	Date of randomization	Censored
No progression and no death	Date of last tumor assessment with no documented progression	Censored
New anticancer treatment started without a prior reported progression	Date of last tumor assessment prior to or on the date of start of new treatment	Censored
Progression documented between scheduled visits	Earliest of: Date of tumor assessment showing new lesion; Date of last tumor assessment (target or non-target) with documented progression.	Progressed
Death without progression and without subsequent anti-cancer therapy and not after >=2 consecutive missed tumor assessment	Date of death	Progressed
Death or progression after two or more consecutive missed tumor assessment	Date of last tumor assessment with documented non-progression	Censored

- PFS accounting for clinical progression (censoring scheme in Table 10.2.2-4). Patients will be defined similarly to the primary definition except that clinical progression will be considered as an event

Table 10.2.2-4. Sensitivity Analysis of PFS Accounting for Clinical Progression

Situation	Date of Progression/Censoring	Outcome
Incomplete or no baseline tumor assessment	Date of randomization	Censored
No adequate post-baseline tumor assessment and no death	Date of randomization	Censored
No progression and no death	Date of last tumor assessment with no documented progression	Censored
New anticancer treatment started without a prior reported progression	Date of last tumor assessment prior to or on the date of start of new treatment	Censored
Progression documented between scheduled visits	Earliest of: Date of tumor assessment showing new lesion; Date of last tumor assessment (target or non-target) with documented progression Scheduled visit that clinical progression is assessed	Progressed
Death without progression and without subsequent anti-cancer therapy	Date of death	Progressed

Listings with pertinent information will be provided.

10.2.3. Duration of Response and Time to Response

Both DOR and time to response will be based on data from independent review, confirmation of CR/PR are not required for these analyses.

DOR is defined as the time from start of the first documentation of objective tumor response (CR or PR) to the first documentation of disease progression (per RECIST v1.1) or death due to any cause, whichever comes first. The analysis of DOR will be performed among the patients who had CR/PR.

DOR will be censored in the following scenario:

- Patients who have not experienced an event (PD or death) will be right-censored on the date of their last evaluable tumor assessment.

DOR = (date of PD or death/censoring – date of objective response + 1)/7 (weeks).

The analysis of DOR will be performed with a Kaplan-Meier method with the same approach as for PFS described in Section 10.2.2. Kaplan-Meier estimates (product-limit estimates) will be presented together with a summary of associated statistics including corresponding two-sided 95% CIs. In particular, the proportion of patients achieved objective response at week 6, 12, 18, 30, and 42 will be estimated with corresponding two-sided 95% CIs. Kaplan-Meier plot of duration of response and corresponding listings of DOR will be provided as well. The sensitivity analysis will be conducted using PP set and tumor response data from investigator's assessment.

Duration of first objective response will be calculated as (date of first objective response – date of randomization +1)/7 in weeks and categorized as 0 to 7, 13, 19, 31, and 43 weeks, cumulative ORR based on first objective response will be calculated within each time period by treatment arm based on ITT population using data from independent reviewer. Difference of ORR and associated 95% CI will be estimated using a generalized linear model adjusted for the randomization stratification factors. Sensitivity analysis will be based on PP population. Subgroup analysis will be based on the subgroup defined in section 6.5.7.

10.2.4. Overall Survival

OS is defined as the time from randomization to date of death due to any cause. For patients who are still alive at the time of data analysis or who are lost to follow up, OS time will be censored at the date of last known to be alive.

OS = (date of death/censoring – date of randomization + 1)/7 (weeks).

The analysis of OS will be performed with a Kaplan-Meier method with the same approach as for PFS described in Section 10.2.2. Kaplan-Meier estimates (product-limit estimates) will be presented together with a summary of associated statistics including corresponding two-sided 95% CIs. In particular, the proportion of overall survival at 18 and 42 weeks will

be estimated with corresponding two-sided 95% CIs. The Kaplan-Meier plot of OS time and listing of OS will provided as well. Sensitivity analysis will be performed based on PP set.

In addition, a Cox proportional regression model will be used to estimate the HR and the corresponding 95% CI for the treatment effect. The main Cox proportional regression model will have treatment and no covariates. Sensitivity analyses will be performed with covariates added to the model, the covariates to be included are stratification factors, race, age, prior radiotherapy, and baseline ECOG performance status.

10.3. Subgroup Analyses

Subgroup analyses of ORR, DCR, PFS, DOR, time to response, and OS will be performed based on the subgroups defined in Section 6.5.7. The subgroup analysis will performed based on ITT set and using data from independent reviewers except for OS. If the endpoint is based on BOR, derived BOR will be used for these analyses. Ratio of ORRs or DCRs, and their associated 90%CI (or difference in ORRs or DCRs, and their associated 95%CI) will be presented using forest plot for subgroup analyses. Number of patients with/without an event, median survival and 95%CI, and survival rate at different time points (18, 30, 42 weeks) will be estimated for subgroup analyses of PFS and DOR. Similarly, those estimates will be provided for OS.

11. Safety Analysis

All analyses of safety will be conducted using the Safety set.

11.1. Adverse Events

All AEs will be coded according to latest available version of MedDRA. The severity of AEs will be graded based on National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) (version 4.03) by investigators. TEAE is defined as any AE that started or deteriorated after the first administration of blinded MYL-1402O or Avastin® through 100 days following the last dose of blinded MYL-1402O or Avastin®. If the onset and end dates for an AE are completely missing or if the onset date is completely missing

and the end date is on or after the first dose of blinded MYL-1402O or Avastin® but prior to or on last dose date plus 100 days, the AE will be considered as treatment-emergent.

AEs will be summarized using the MedDRA PT as event category and/or MedDRA primary SOC as summary category. All AE tables will be restricted to treatment-emergent AEs unless otherwise specified.

Each patient will be counted only once within each PT or SOC. If a patient experiences more than one AE within a PT or SOC, only the AE with the strongest relationship or the worst severity, as appropriate, will be included in the summaries of relationship and severity. AEs with missing classifications regarding relationship to study treatment and start date greater or equal to start of IMP will be considered as related to IMP. AEs with missing toxicity grade will be counted into 'any grade' in the summarization by toxicity grade.

TEAEs will be sorted in descending order of frequency (total column) by SOC and by PT within each SOC, unless otherwise stated, alphabetical order will be used for same frequency of SOC or PT. The dictionary version used for reporting the study will be described in the relevant table and listing footnotes.

11.1.1. All Adverse Events

An overview summary of the number and percentage of patients with any TEAE, serious TEAE, IMP-related TEAE, IMP-related serious TEAE, TEAE leading to IMP discontinuation, IMP-related TEAE leading to IMP discontinuation, TEAE leading to death, and IMP-related TEAE leading to death will be provided by treatment arm and overall. If an AE is associated with toxicity of 5 or outcome of death, the AE will be included into 'TEAE leading to death' summary.

Summaries of the total number of TEAEs and the number and percentage of patients with at least one TEAE will be provided. The number and percentage of patients and the number of events will also be presented by SOC and PT within each treatment arm and overall. Listing of AEs including all relevant information such as AE SOC/PT, start/stop date, duration of AE, toxicity grade, relationship to IMP/CP, action taken with IMP/CP, and outcome etc.,

will be provided. AEs started prior to the first dose of IMP or after last dose of IMP + 100 days will be identified in the listing.

11.1.2. Relationship of Adverse Events to IMP

A summary of TEAEs related to IMP will be presented in a table using frequency of occurrence by SOC and PT. The investigator will provide an assessment of the relationship of the event to the IMP. The possible relationships are "Not Related", "Unlikely", "Possibly", "Probably", and "Definitely". 'Possibly', 'Probably', and 'Definitely' will be considered as 'Related' while 'Not Related' and 'Unlikely' will be considered as 'Not Related' for the analysis. If a patient reports multiple occurrences of the same TEAE, only the most closely related occurrence will be presented. TEAEs that are missing a relationship will be presented in the summary table as "Related" but will be presented in the data listing with a missing relationship. Percentages will be calculated out of the number of patients in the Safety Set.

11.1.3. Severity of Adverse Events

A summary of TEAEs by worst (i.e. highest) CTCAE grade will be presented in a table by SOC and PT for each treatment arm and overall. The possible CTCAE grades are 1 = Mild, 2 = Moderate, 3 = Severe, 4 = Life-threatening, and 5 = Death, they will be summarized as 'Any grade', 'Grade 3-4', and 'Grade 5'. If a patient reported multiple occurrences of the same TEAE, only the highest grade captured on the eCRF will be presented for a given SOC/PT and grade group. TEAEs that are missing severity will be included into 'Any grade'. Percentages will be calculated out of the number of patients in the Safety Set.

11.1.4. Serious Adverse Events

The seriousness of an AE should be assessed by the Investigator independently from the severity of the AE. A serious AE (SAE) is defined as any untoward medical occurrence that at any dose results in death, is life-threatening, results in persistent or significant disability/incapacity, is a congenital anomaly, or requires inpatient hospitalization or prolongation.

Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered SAEs when, based upon medical judgment, they may jeopardize the patient and may require medical or surgical intervention to prevent one of the outcomes listed above.

The serious TEAEs will be categorized and presented by SOC and PT as well as by combined relationship to IMP by treatment arm and overall. Listing of SAEs including all relevant information such as AE SOC/PT, start/stop date, duration of AE, toxicity grade, relationship to IMP/CP, action taken with IMP/CP, seriousness, and outcome etc., will be provided. SAEs started prior to the first dose of IMP or after last dose of IMP + 100 days will be identified in the listing.

11.1.5. Adverse Events Leading to IMP Discontinuation

'Drug Withdrawn' with IMP will be collected on the eCRF on which a summary of TEAEs leading to IMP discontinuation will be based. This summary will include the number and percentage of patients by SOC and PT within each treatment arm and overall. Related TEAEs leading to IMP discontinuation will also be summarized by SOC and PT using count and percentage. Listing of AEs leading to IMP discontinuation including all relevant information such as AE SOC/PT, start/stop date, duration of AE, toxicity grade, relationship to IMP/CP, action taken with IMP/CP, and outcome etc., will be provided.

11.1.6. Adverse Events of Interest

Adverse events will be searched against Standardized MedDRA Queries (SMQs) and mapped to SMQ terms and scopes (broad, narrow), the following SMQs will be analyzed in the study. AEs, SAEs, and grade >= 3 AEs will be summarized by SMQ/preferred term, and sorted in descending order of frequency (total column) by SMQ and by PT within each SMQ, alphabetical order will be used for same frequency of SMQ or PT.

- Hypersensitivity (narrow)
- Hypertension (narrow)
- Proteinuria (broad),

- Embolic and thrombotic events, vessel type unspecified and mixed arterial and venous (narrow)
- Haemorrhage terms (excl laboratory terms) (narrow)
- Interstitial lung disease (broad)
- Embolic and thrombotic events, venous (narrow)
- Cardiac failure (broad)
- Gastrointestinal perforation (narrow)
- Osteonecrosis (broad)

11.1.7. Death

All deaths will be tabulated and listed for patients in the Safety set. The death table will include the following information:

- Number of patients who died
  - Primary reason for death
    - Disease progression
    - Adverse event
    - Other
    - Unknown
- Number of patients who died within 30 days of the last IMP administration
  - Primary reason for death
    - Disease progression
    - Adverse event
    - Other
    - Unknown

The listing of deaths will be provided with all the relevant information such as death date and reason for death. The death data will be ascertained from the dedicated Death eCRF form.

11.2. Clinical Laboratory Evaluations

Laboratory abnormalities are classified according to NCI-CTCAE version 4.03 or based on normal ranges collected from laboratories. Safety laboratory assessment will be performed locally, test values and normal ranges will be converted based on International System of Units prior to toxicity grading. The toxicity grading is only related to the lab values itself and does not respect the non-numeric information as described in the CTCAE grading definition (Appendix I). CTCAE gradable parameters and associated toxicities are listed in Table 11.2-1.

Table 11.2-1. CTCAE Gradable Parameters and Associated Toxicities and Grades

Panel	Parameter (Unit)	Low direction toxicity (Grade)	High direction toxicity (Grade)
Hematology	Hemoglobin (HB) (g/L)	Anemia (1, 2, 3)	
	Leukocytes (10e9/L)	white blood cell decreased (1, 2, 3, 4)	
	Neutrophils/ absolute neutrophils count (ANC) (10e9/L)	neutrophil count decreased (1, 2, 3, 4)	
	Platelet count (PLT) (10e9/L)	platelet count decreased (1, 2, 3, 4)	
Chemistry	Albumin (g/L)	Hypoalbuminemia (1, 2, 3)	
	Alkaline phosphatase (ALP) (U/L)		alkaline phosphatase increased (1, 2, 3, 4)
	Alanine aminotransferase (ALT) (U/L)		ALT increased (1, 2, 3, 4)
	Aspartate aminotransferase (AST) (U/L)		AST increased (1, 2, 3, 4)
	Bilirubin (total) (µmol/L)		blood bilirubin increased (1, 2, 3, 4)
	Creatinine (µmol/L)		creatinine increased (1, 2, 3, 4)
	Potassium (mmol/L)	Hypokalemia (2, 3, 4)	Hyperkalemia (1, 2, 3, 4)
	Sodium ((mmol/L)	Hyponatremia (1, 3, 4)	Hypernatremia (1, 2, 3, 4)

Coagulation	Prothrombin time INR	INR increased (1, 2, 3)
	Activated Partial thromboplastin time (aPTT) (second)	aPTT prolonged (1, 2, 3)

For these parameters which are graded with two toxicities such as potassium (hypokalemia/ hyperkalemia), the toxicities will be summarized separately. Low direction toxicity (eg. hypokalemia) grades at baseline and post-baseline will be set to 0 when the variables are derived for summarizing high direction toxicity (e.g. hyperkalemia), and vice versa.

For non-CTCAE gradable parameters collected in this study, their abnormalities are assessed as low, high, normal based on the comparison of observed values with normal ranges.

11.2.1. Serum Chemistry, Hematology, and Coagulation

For parameters (chemistry, hematology, coagulation) with numerical results reported, summary statistics for observed values and changes from baseline will be tabulated at each scheduled time point by treatment arm and overall for patients in the Safety Set.

CTCAE gradable parameters

The laboratory toxicities (chemistry, hematology, coagulation, Table 11.2-1) will be tabulated by the shift of CTCAE grade from baseline to worst post-baseline grade using descriptive statistics (count and percentage). The highest post-baseline CTCAE grade is considered as the worst grade for the summary.

- The shift table will summarize baseline CTCAE grade vs. the worst post-baseline CTCAE grade (grade = 0, 1, 2, 3, 4) for each parameter/toxicity, patients with a baseline value and at least 1 post-baseline values will be included for the analysis.

Non-CTCAE gradable parameters

Most tests in hematology and chemistry panels are CTCAE gradable in this study, summary of shift from baseline (low, normal, high) to minimum/maximum post-baseline value is not required.

The listings (hematology, chemistry, coagulation) will include all the laboratory parameters as available in the database with the relevant information such as visit, assessment date,

parameter, value, normal ranges etc. Listings will be sorted by treatment arm, patient identifier, parameter, assessment date or visit.

Liver function parameters

ALT, AST, ALP, and total bilirubin are used individually or together to assess possible drug induced liver toxicity. The ratios of test result over upper limit of normal (ULN) for individual test or combined tests will be calculated and classified for these parameters during the post-baseline period.

Summary of liver function tests will include the following categories. The number and percentage of patients with at least one occurrence for each of the following categories will be summarized by treatment arm and overall, if applicable:

- AST ≥3*ULN / ≥5*ULN / ≥10*ULN / ≥20*ULN.
- ALT ≥3*ULN / ≥5*ULN / ≥10*ULN / ≥20*ULN.
- (ALT or AST) ≥ 3*ULN/ ≥ 5*ULN/ ≥ 10*ULN/ ≥ 20*ULN
- Total bilirubin ≥2*ULN
- Concurrent ALT ≥ 3*ULN and TBILI ≥ 2*ULN
- Concurrent AST ≥ 3*ULN and TBILI ≥ 2*ULN
- (ALT or AST) ≥3*ULN concurrently with total bilirubin ≥2*ULN.
- (ALT or AST) ≥3*ULN concurrently with total bilirubin ≥2*ULN and ALP >2*ULN.
- (ALT or AST) ≥3*ULN concurrently with total bilirubin ≥2*ULN and (ALP ≤2*ULN or missing).

Concurrent measurements are those occurring on the same date.

Categories will be cumulative, i.e a patient with an elevation of AST ≥10*ULN will also appear in the categories ≥5*ULN and ≥3*ULN. Liver function elevation and possible Hy's Law cases will be summarized using frequency and percentage by treatment arm.

Listing of patients with ALT or AST ≥3*ULN or total bilirubin ≥2*ULN will include variables patient identifier, visit, date of collection, study day, parameter (ALT, AST, ALP, total bilirubin), result, unit, result/ULN (ratio of result over ULN), CTCAE grade. This listing will be sorted by treatment arm, patient identifier, visit/ data of collection, study day, and parameter.

11.2.2. Urinalysis and Pregnancy Test

Urinalysis data and pregnancy tests will be presented in separate listings, which will be sorted by treatment arm, patient identifier, parameter, assessment date or visit.

11.3. Vital Sign Measurements

A summary table presenting observed values and changes from baseline will be presented for vital sign data, including systolic blood pressure (mmHg), diastolic blood pressure (mmHg), pulse (beats/min), and weight (kg), by treatment arm and overall for patients in the Safety Set. Changes from baseline to each scheduled post-baseline visit will be presented.

All vital sign data by patient will be presented in a listing for patients in the Safety set.

11.4. ECOG Performance Status

The ECOG performance status score will be presented at each scheduled visit to show the number and percentage of patients in each category for patients in the Safety Set. The ECOG shift from baseline to highest post-baseline score will be summarized by treatment arm and overall based on Safety set, patients with a baseline value and at least 1 post-baseline value will be included into the analysis.

All ECOG performance status scores along with relevant information will be presented in a listing for patients in the Safety Set.

12. Pharmacokinetic Analysis

Details regarding the analysis of pharmacokinetic (PK) data will be provided in a separate analysis plan.

13. Immunogenicity Analysis

The immunogenicity or antidrug antibodies (ADA) samples will be collected at Baseline (pre-dose), Cycle 2, 4, and 6 in Period 1, Cycle 704 and 708 in Period 2, and at Safety Follow-up visit. Treatment emergent ADA includes treatment induced ADA and treatment boosted ADA. If patients do not have positive ADA results prior to treatment but have at least one positive post-baseline ADA result, these patients are categorized as having treatment induced ADA; if patients have positive ADA results prior to treatment and the titer $\geq$ 4*baseline titer while on treatment, these patients are categorized as having treatment boosted ADA. Immunogenicity data will be summarized as follows for Week 18 and 42 analyses unless otherwise specified, the analyses will be based on safety set unless otherwise specified:

- Summary of immunogenicity titer results by visit and treatment;
- Summary of ADA samples analyzed by visit and treatment;
- Summary of NAb (neutralizing antibody) samples analyzed by visit and treatment;
- Summary of immunogenicity results by visit and treatment excluding baseline positive patients;
- Overall incidence of treatment emergent antidrug antibodies during the study;
- Overall summary of immunogenicity for Week 42 analysis;
- Summary and analysis of overall response rate (ORR) based on BOR derived from independent reviewer data for patients with at least one positive ADA assessment and for patients without any positive ADA assessment for the ITT set;
- Summary of hypersensitivity (SMQ) adverse events by ADA status. The ADA status is categorized as ADA-negative post-baseline, which contains patients with negative ADA results at baseline and post-baseline, and ADA-positive post-baseline, which contains patients with at least one positive post-baseline ADA results and any baseline ADA results (positive, negative or missing).

The corresponding listing including relevant information will be provided as well.

14. Changes to the Planned Analyses in the Protocol

- First dose date is defined as Day 1 for the analysis to be compliant with CDISC requirement.
- To evaluate the difference in the ORRs at Week 18, an asymptotic 2-sided 95% CI for the difference in ORRs at Week 18 will be calculated using Wald confidence interval (OPTION=RISKDIFF in TABLES statement in PROC FREQ) without adjustment for covariates instead of using an unstratified Cochran-Mantel-Haenszel (CMH) test as specified in Protocol Section 7.7.1.2. The method will be used for the evaluation of difference in DCR.
- MITT set was added for BSSR analysis.

15. References

Blackwelder WC. Sample size and power for prospective analysis of relative risk. Statistics in Medicine 1993; 12:691-698.

Brookmeyer R, Crowley J. A Confidence Interval for the Median Survival Time. Biometrics 1982; 38:29–41.

Eisenhauer EA. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer 2009; 45:228-47.

Ellis LM, Bernstein DS, Voest EE, Berlin JD, Sargent D, Cortazar P, Garrett-Mayer E, Herbst RS, Lilenbaum RC, Sima C, Venook AP, Gonen M, Schilsky RL, Meropol NJ, Schnipper LE. American Society of Clinical Oncology perspective: Raising the bar for clinical trials by defining clinically meaningful outcomes. J Clin Oncol. 2014 Apr 20; 32(12):1277-80.

He K, Chen H, Gwise T. Statistical Considerations in Evaluating a Biosimilar Product in a Oncology Clinical Study. Clin cancer Res. 2016; 22: 5167-5170

Jennison C, Turnbull BW. In: Group Sequential Methods with Applications to Clinical Trials. Chapman & Hall/CRC, Boca Raton, 2000.

Johnson DH, Fehrenbacher L, Novotny WF, et al. Randomized phase II trial comparing bevacizumab plus carboplatin and paclitaxel with carboplatin and paclitaxel alone in previously untreated locally advanced or metastatic non-small-cell lung cancer. J Clin Oncol. 2004; 22(11):2184-91.

Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons 1980.

Niho S, Kunitoh H, Nokihara H, et al;JO19907 Study Group. Randomized phase II study of first-line carboplatin-paclitaxel with or without bevacizumab in Japanese patients with advanced non-squamous non-small-cell lung cancer. Lung Cancer. 2012; 76(3):362-7.

Reck M, von Pawel J, Zatloukal P, et al; BO17704 Study Group. Overall survival with cisplatin-gemcitabine and bevacizumab or placebo as first-line therapy for nonsquamous non-small-cell lung cancer: results from a randomised phase III trial (AVAiL). Ann Oncol. 2010; 21(9):1804-9.

Sandler A, Gray R, Perry MC, et al. Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. N Engl J Med. 2006; 355(24):2542-50.

Zhou C, Wu YL, Chen G, et al. BEYOND: A randomized, double-blind, placebo-controlled, multicenter, phase III study of first-line carboplatin/paclitaxel plus bevacizumab or placebo in Chinese patients with advanced or recurrent nonsquamous non–small-cell lung cancer. J Clin Oncol. 2015; 33(19):2197-2204.

16. Appendices

Appendix I. CTCAE Grading Criteria for Safety Laboratory Assessments

Adverse Event	Grade 1	Grade 2	Grade 3	Grade 4
Anemia	HB <LLN – 100 g/L	HB <100 – 80 g/L	HB <80 g/L	

White blood cell decrease	<LLN – 3.0 × 10e9/L	<3.0 – 2.0 × 10e9/L	<2.0 – 1.0 × 10e9/L	<1.0 × 10e9/L
Neutrophil count decreased	<LLN – 1.5 × 10e9/L	<1.5 – 1.0 × 10e9/L	<1.0 – 0.5 × 10e9/L	<0.5 × 10e9/L
Platelet count decreased	<LLN – 75.0 × 10e9/L	<75.0 – 50.0 × 10e9/L	<50.0 – 25.0 × 10e9/L	<25.0 × 10e9/L
Hypoalbuminemia	<LLN – 30 g/L	<30 – 20 g/L	< 20 g/L	
Alkaline phosphatase increased	>ULN – 2.5ULN	>2.5ULN – 5.0ULN	>5.0ULN – 20.0ULN	>20.0ULN
ALT increased	>ULN – 3.0ULN	>3.0ULN - 5.0ULN	>5.0ULN – 20.0ULN	>20.0ULN
AST increased	>ULN – 3.0ULN	>3.0ULN - 5.0ULN	>5.0ULN – 20.0ULN	>20.0ULN
Blood bilirubin increased	>ULN – 1.5ULN	>1.5ULN - 3.0ULN	>3.0ULN – 10.0ULN	>10.0ULN
Creatinine increased	>ULN – 1.5ULN	>1.5ULN - 3.0ULN	>3.0ULN – 6.0ULN	>6.0ULN
Hypokalemia		<LLN – 3.0 mmol/L	<3.0 – 2.5 mmol/L	<2.5 mmol/L
Hyperkalemia	>ULN – 5.5 mmol/L	>5.5 – 6.0 mmol/L	>6.0 – 7.0 mmol/L	>7.0 mmol/L
Hyponatremia	<LLN – 130 mmol/L		<130 – 120 mmol/L	<120 mmol/L
Hypernatremia	>ULN – 150 mmol/L	>150 - 155 mmol/L	>155 – 160 mmol/L	>160 mmol/L
aPTT prolonged	>ULN – 1.5ULN	>1.5ULN – 2.5ULN	>2.5ULN	
INR increased	>ULN - 1.5ULN	>1.5ULN – 2.5ULN	>2.5ULN	

Appendix II. Data Cut-Off Rules for the Analyses

- Data cut off will be performed at raw data level prior to creating SDTM datasets.
- All data will be included for main CSR and final CSR analyses at week 18 and 42, respectively, for screening failed patients and not dosed patients.
- For week 18 analysis, BOR assessment should not be missing unless there is no post-baseline tumor assessment due to early termination of IMP or baseline failure/not dosed, the maximum of Week 18 visit date and BOR assessment date will be used as data cut-

off date if both dates are not missing; if Week 18 visit date is missing, the maximum of first dose date + 129 and BOR assessment date will be used as data cut-off date.

- Week 42 visit date will be used as data cut-off date for patients who are treated or on-study beyond week 42.
- First dose date + 297 will be used as data cut-off date for patients who terminated early for final CSR analyses at week 42.
- Data assessment/collection/event date will be compared with data cut-off date for each individual patient, data after cut-off date will be excluded for the analysis.
- For AEs or concomitant medications, if their start dates are on or prior to date cut-off dates, but end dates are after data cut-off dates, end dates will be set to missing, ongoing flags will be set to 'Yes', AE outcomes will be set to 'Not Recovered/Not Resolved'.