Document Type:	Statistical Analysis Plan
Official Title:	An Open-label, Single-arm, Phase II Study of Regorafenib and Nivolumab in Patients with Mismatch Repair-Proficient (pMMR)/Microsatellite Stable (MSS) Colorectal Cancer (CRC)
NCT Number:	NCT04126733
Document Date:	11 NOV 2020

Title page

An Open-label, Single-arm, Phase II Study of Regorafenib and Nivolumab in Patients with Mismatch Repair-Proficient (pMMR)/Microsatellite Stable (MSS) Colorectal Cancer (CRC)

Regorafenib and Nivolumab in pMMR/MSS CRC

Bayer study drug	BAY 73-4506		
Study purpose:	Determine the safety and efficacy of the combination of regorafenib and nivolumab		
Clinical study phase:	II	**Date:**	11 NOV 2020
Study No.:	20975	**Version:**	2.0
Author:	PPD CCI		

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.

The approval of the Statistical Analysis Plan is documented in a separate Signature Document.

Table of Contents

Title page **1**

Abbreviations **3**

1. Introduction **4**

2. Study Objectives **4**

3. Study Design **4**

4. General Statistical Considerations **4**
4.1 General Principles 4
4.2 Handling of Dropouts 4
4.3 Handling of Missing Data 5
4.4 Interim Analyses and Data Monitoring 5
4.5 Analysis of long-term follow-up 5
4.6 Data Rules 6
4.7 Validity Review 6
4.8 Subgroup analyses 6

5. Analysis Sets **7**
5.1 Assignment of analysis sets 7

6. Statistical Methodology **7**
6.1 Population characteristics 8
6.2 Efficacy 8
6.2.1 Sample size 8
6.2.2 Primary Endpoint: Overall response rate (ORR) based on RECIST v1.1 9
6.2.3 Secondary endpoints 10
CCI [redacted]
6.5 Safety 11
6.5.1 Extent of exposure 12
6.5.2 Adverse events 12
6.5.3 Deaths and Serious Adverse events 14
6.5.4 Pregnancies 14
6.5.5 Clinical laboratory data 14
6.5.6 12-Lead ECG, QTc 16
6.5.7 ECOG status 17
6.5.8 Other safety measures 17

7. Document history and changes in the planned statistical analysis **17**
CCI [redacted]

Abbreviations

AE	Adverse event
Anti-PD-1	Anti-programmed cell death protein-1
BICR	Blinded independent central review
BMI	Body mass index
CI	Confidence interval
CRF	Case report form
CRC	Colorectal cancer
CSR	Clinical study report
CTCAE	Common terminology criteria adverse event
ECG	Electrocardiogram
FAS	Full analysis set
IMAE	Immune modulating adverse event
MedDRA	Medical dictionary for regulatory activities
MSS	Microsatellite stable CRC
N/A	Not applicable
NCI	National cancer institute
PD	Pharmacodynamics
pMMR	mismatch repair-proficient CRC
PK	Pharmacokinetics
QT	Q wave to T wave interval
QTc	Corrected QT
RECIST	Response evaluation criteria in solid tumors
SAE	Serious adverse event
SAP	Statistical analysis plan
TEAE	Treatment emergent AE
VEGF	Vascular endothelial growth factor
WHO	World health organization

CCI

[redacted]

2. Study Objectives

The purpose of this study is to determine the safety and efficacy of the combination of regorafenib and nivolumab in patients with advanced or metastatic pMMR/MSS colorectal cancer.

3. Study Design

This is an open-label single-arm study.

4. General Statistical Considerations

4.1 General Principles

The statistical evaluation will be performed by using the software package SAS release 9.2 or higher (SAS Institute Inc., Cary, NC, USA). All variables will be analyzed by descriptive statistical methods. The number of data available and missing data, mean, standard deviation, minimum, quartiles, median, and maximum will be calculated for metric data. Frequency tables will be generated for categorical data.

All data from the data base will also be given in subject data listings. Those will not be described in further detail as they will just represent the database.

4.2 Handling of Dropouts

All patients who discontinue the study after enrollment for any reason except death before the protocol defined end are considered drop-outs. Data of drop-outs will be used up to the point of study discontinuation, however no imputation for missing values will be done, except for dates (see below).

4.3 Handling of Missing Data

All missing or partial data will be presented in the subject data listing as they are recorded on the Case Report Form (CRF).

The number of subjects who prematurely discontinue the study and study treatment for any reason, as well as the reasons for premature discontinuation of study and study treatment, will be reported.

Missing or unevaluable tumor assessments, including scheduled assessments that were not done and incomplete assessments that did not result in an unambiguous tumor response evaluation according to RECIST 1.1, will not be used in the calculation of derived efficacy variables related to tumor assessments unless a new lesion occurred or the lesions that were evaluated already showed progressive disease. No imputation will be performed for missing lesion assessments and tumor response evaluation.

If a date is incomplete, (e.g. only year and month of date of tumor assessment or date of death is available), then day 15 of the month will be used for the calculation of, for example, overall survival and progression free survival. For imputation of AE dates, if the range of possible dates (e.g. a month, if day is missing) and during this time period “treatment emergent” and “non treatment emergent” is possible, the date will be imputed such that the AE will be “treatment emergent” (e.g. the day of treatment start if day is missing).

If the actual scan date of the radiological progression is missing and radiological or clinical progression has been documented based on criteria specified in the protocol, the scheduled scan date will be used to calculate the time to progression.

Safety variables, medical history and concomitant medications:

Treatment emergent AEs, treatment phases, period and relative days will be derived according to data management programming, the operations manual and standard guidelines.

4.4 Interim Analyses and Data Monitoring

No formal interim analysis is planned in this study.

However it is planned to look at the data after 30-40 participants have had at least the first response evaluation, in order to get better information for the planning of the intended Phase III study. This data look will be focused on the response data only and the analyses will be described in a separate data look SAP.

No formal conclusions for this study will be drawn from this look at the data and due to the exploratory nature of this study, no formal adjustment of the type-I error will be performed.

4.5 Analysis of long-term follow-up

During the long term follow-up data on survival and subsequent anti-cancer therapies are collected by either telephone contact or onsite visit. After the last patient finished the long-term follow-up, a final analysis of the long-term follow-up will be conducted. This will be conducted on the follow-up set of patients and will include population characteristics, overall survival analysis and anti-cancer medication. Details of the FU analysis will be specified in a separate document, either as addendum to this SAP or in a separate SAP.

4.6 Data Rules

Dates:

Generally, for each date stored in the database a set of organizational variables will be derived in order to describe the temporal context of that date in the specific study in terms of: the phase of treatment (pre, during or post study treatment), the day relative to the start of study treatment, the day relative to the end of study treatment.

Year will be calculated as (days/365.25) rounded up to 1 decimal place.

Month will be calculated as (days/30.44) rounded up to 1 decimal place.

Time to event or duration of event endpoints will be based on the actual date rather than Visit Day.

Baseline:

Unless specified otherwise, baseline is the last pre-dose measurement performed prior to the first administration of the first study drug.

Post-Baseline:

Every value taken after the Baseline measurement is considered a post-Baseline value.

Change from Baseline:

Change from baseline is defined as (post-baseline value - baseline value).

Indirect Bilirubin:

Indirect Bilirubin is calculated as:

total Bilirubin – direct Bilirubin

Indirect Bilirubin was entered in the CRF at the beginning of the study. These values will not be analyzed and only presented in the subject data listings for completeness. Summary statistics will only be created from the derived indirect Bilirubin as per formula given above.

Unscheduled Visits:

Unscheduled Visits will not be used in the analysis (i.e. section 14) and will only be presented as part of the subject data listing (i.e. section 16).

4.7 Validity Review

The results of the final validity review meeting will be documented in the final list of important deviations, validity findings and assignment to analysis set(s). Any changes to the statistical analysis prompted by the results of the review of study data will be documented in an amendment and, if applicable, in a supplement to this SAP.

4.8 Subgroup analyses

All efficacy and adverse events analyses are also done by subgroups for the following subgroups:

- Sex
- Race
- Age (<65, 65-74, 75-84, >=85)

- Presence of liver metastases at baseline (yes / no)

  The presence of liver metastases is defined as having a liver lesion (target or non-target) at baseline. The respective question of the baseline cancer characteristics page is not used for this subgroup definition.

- Primary cancer (colon / rectal)

5. Analysis Sets

5.1 Assignment of analysis sets

Final decisions regarding the assignment of subjects to analysis sets will be made during the validity review of study data and documented in the final list of important deviations, validity findings and assignment to analysis set(s) (see section 4.7).

Population	Description
Enrolled	All participants who sign the ICF
FAS	All participants who have received at least 1 dose of study intervention will be used for efficacy analysis.
SAF	All participants who have received at least 1 dose of study intervention will be used for safety analysis.
IMS	All patients who have received at least one dose of nivolumab and have at least one ADA sample taken (during the treatment or follow-up observation period) that is appropriate for ADA testing (with reportable result) will be included in the analysis of the immunogenicity of nivolumab, which is defined as the immunogenicity analysis set.
Follow-Up (FUP)	All patients who entered the long-term follow-up and have had at least one data collection during the long-term follow-up will be analyzed in the final analysis after study end.

Abbreviations: ADA = anti-drug antibodies; FAS = full analyses set; ICF = informed consent form; PK = pharmacokinetics; r-PKS = regorafenib PK analysis set; n-PKS = nivolumab PK analysis set; IMS = immunogenicity analysis set

No additional treatment assignments to compensate for potential unevaluable patients (e.g. lack of post baseline tumor scans) beyond the needed number of 70 treated patients is planned. Therefore SAF and FAS will be equal and only FAS will be referenced in the analyses.

6. Statistical Methodology

All variables will be analyzed at least descriptively, which includes summary statistics (N, mean, std. dev. minimum, median, maximum) for continuous variables and frequency tables (N, % per category) for categorical variables.

A confirmatory analyses will be conducted only for the primary endpoint at one-sided type-I error level of 2.5% or two-sided type-I error level of 5% respectively. No further confirmatory analysis is intended, and no further tests are conducted. Confidence intervals will be presented for selected variables.

Covid-19:

All COVID-19 related findings will be presented as a separate listing and in applicable tables (e.g. protocol deviations) as separate finding as well.

6.1 Population characteristics

Baseline values and demographic characteristics are defined in section 8.2.7 of the protocol.

All baseline and demographic values will be summarized by descriptive tables as described above.

Baseline cancer characteristics:

Baseline cancer characteristics will be summarized by a table using appropriate summaries. The lines of previous therapies will be analyzed as continuous variable, as well as categorical variable. The variable "liver metastases at baseline according to scans" which is used for the subgrouping of tables, will be added to this table as well as information about patients who had a liver resection.

Medical history

Data for medical history will be coded using the most recent version of MedDRA. Summary statistics (frequency and percentage) will be provided by system organ class (SOC), high level term (HLT) and preferred term (PT).

Medications

The dictionary used for coding medications is the WHO Drug dictionary. Systemic anti-cancer therapies will be displayed by frequency of subjects for each drug, (prior and during follow-up). Diagnostic and therapeutic procedures will be tabulated by frequency of subjects by procedure (prior, concurrent and during follow-up), other concomitant medications will be presented by frequency of subjects for each drug category.

In addition the prior anti-cancer therapies will be displayed, grouped by:

1. Neoadjuvant/adjuvant setting
2. Advanced/metastatic setting

Disposition of Subjects

Since the study is only conducted in the US, the number of subjects enrolled and included in each of the analysis populations will be tabulated overall and by center only. A summary table will also be presented for the number of subjects enrolled and the number and percentage of subjects in each of the defined populations. The reasons for subjects excluded from each of the subject populations will also be tabulated. The reasons for discontinuation of study treatment will be tabulated.

6.2 Efficacy

Analyses will be based on the FAS population.

6.2.1 Sample size

As of amendment 2, it was decided to stop recruitment at 70 treated patients. With this sample size of 70 patients, when the true underlying ORR is 17% or higher and a baseline response rate of 5% is assumed, a one-sided exact binomial test at a type-I error of at approximately 2.5% would result in a power of 92.6%, with at least 8 responders needed to achieve significance.

The location of the primary tumor is a factor and it is expected that 40%-45% of patients will have the primary tumor on the right side. With 31 patients with right-sided tumors, the probability of observing 4 or more responders (ORR 13%) is approximately 80% when the

true underlying response rate is 17%, and the probability of observing 2 or fewer responders (ORR 6.5%) is approximately 80% when the true ORR is 5%.

With 39 patients with left-sided tumors, the probability of observing 6 or more responders (ORR 15%) is approximately 67% when the true underlying response rate is 17%, and the probability of observing 2 or fewer responders (ORR 5.1%) is approximately 70% when the true ORR is 5%.

For this calculation SAS PROC POWER was used.

6.2.2 Primary Endpoint: Overall response rate (ORR) based on RECIST v1.1

Primary endpoint for efficacy assessment is ORR per RECIST v1.1. Radiological tumor assessments will be done locally by the investigator.. ORR is defined as the percentage of subjects with complete response (CR) or partial response (PR) and will be analyzed after all participants met the criteria for primary completion, which is to have 5 post-baseline scans i.e., to have been followed for 10 months, or to have dropped out of the trial. The data base will then be locked for analysis.

The primary analysis of the primary endpoint will investigate the hypothesis

H_0: $ORR_{Combination} \leq 5\%$ versus

H_1: $ORR_{Combination} > 5\%$

The hypothesis will be tested by a one-sided exact binomial test, assuming a background response rate for the combination to be at most 5%. The target response rate for the combination treatment will be 17%. Using a one-sided exact binomial test at a type-I error of at most 2.5% at least 8 responders out of the 70 patients are needed to achieve significance.

The result of the exact binomial test will be presented in a table, giving the response rate, the number of responders needed, the number of responders observed, the test-statistic and the p-value. Inferential testing will not be repeated by subgroup, only confidence intervals will be presented for the subgroup analysis.

Further analysis of the primary endpoint includes frequency tables (with the ORR table also including all “Best overall response” categories) as well as corresponding 95% two-sided Clopper-Pearson confidence intervals. RECIST results will be displayed by visit as well. All descriptive analyses (incl. confidence intervals) of the primary endpoint will be repeated by two subgroups:

- primary tumor side (right and left). The information about tumor side does not come from the CRF, but is stored in the IxRS system and has to be transferred to the database. .
- ECOG performance status at baseline (0 vs 1)

Formal testing of ORR will not be done for these subgroups.

6.2.3 Secondary endpoints

Overall survival (OS):

Overall survival is defined as the time (days) from the date of first treatment to the date of death.

OS for subjects who are still alive and in the study or who are lost to follow-up at the point of analysis will be censored at the last available date.

OS for subjects who have no post baseline information will be censored at day 1.

Progression free survival (PFS):

PFS is defined as the time (days) from the date of first treatment to the date of radiological or clinical disease progression or death due to any cause, if death occurs before progression is documented.

PFS for subjects without radiological or clinical tumor progression or death at the time of analysis will be censored at their last date of tumor evaluation.

PFS for subjects who have no tumor assessment post baseline and are still alive will be censored at day 1. Patients that report a PD after missed tumor assessment(s) will be censored at the last scan date the patient was known to be progression free.

Disease control rate (DCR):

Disease control rate is defined as the proportion of patients whose best response was complete response (CR), partial response (PR) or stable disease (SD). DCR will be evaluated at 8, 16, 24, 32 and 40 weeks

Scans collected after first progression disease (PD) will not be considered.

Duration of response (DOR):

Duration of response (for PR and CR) is defined as the time (days) from the first documented objective response of PR or CR, whichever is noted earlier, to disease progression or death (if death occurs before progression is documented). Duration of response will be defined for responders only, i.e. subjects with CR or PR.

Time to response (TTR):

Time to response is defined as the time (days) from first treatment until CR or PR and will be derived for responders only.

Duration of stable disease (DoSD):

Duration of stable disease is defined as the time (days) from first treatment to date that disease progression or death (if death occurs before progression) is first documented. This variable will only be calculated for subjects who failed to achieve a best response of CR or PR (i.e. calculated for subjects with SD as best overall response). Also subjects, who have no evaluable tumor assessments post-baseline available, will be excluded. The actual dates that the tumor scans were performed will be used for this calculation. Duration of stable disease for subjects without disease progression or death before progression at the time of analysis will be censored at the date of their last tumor assessment.

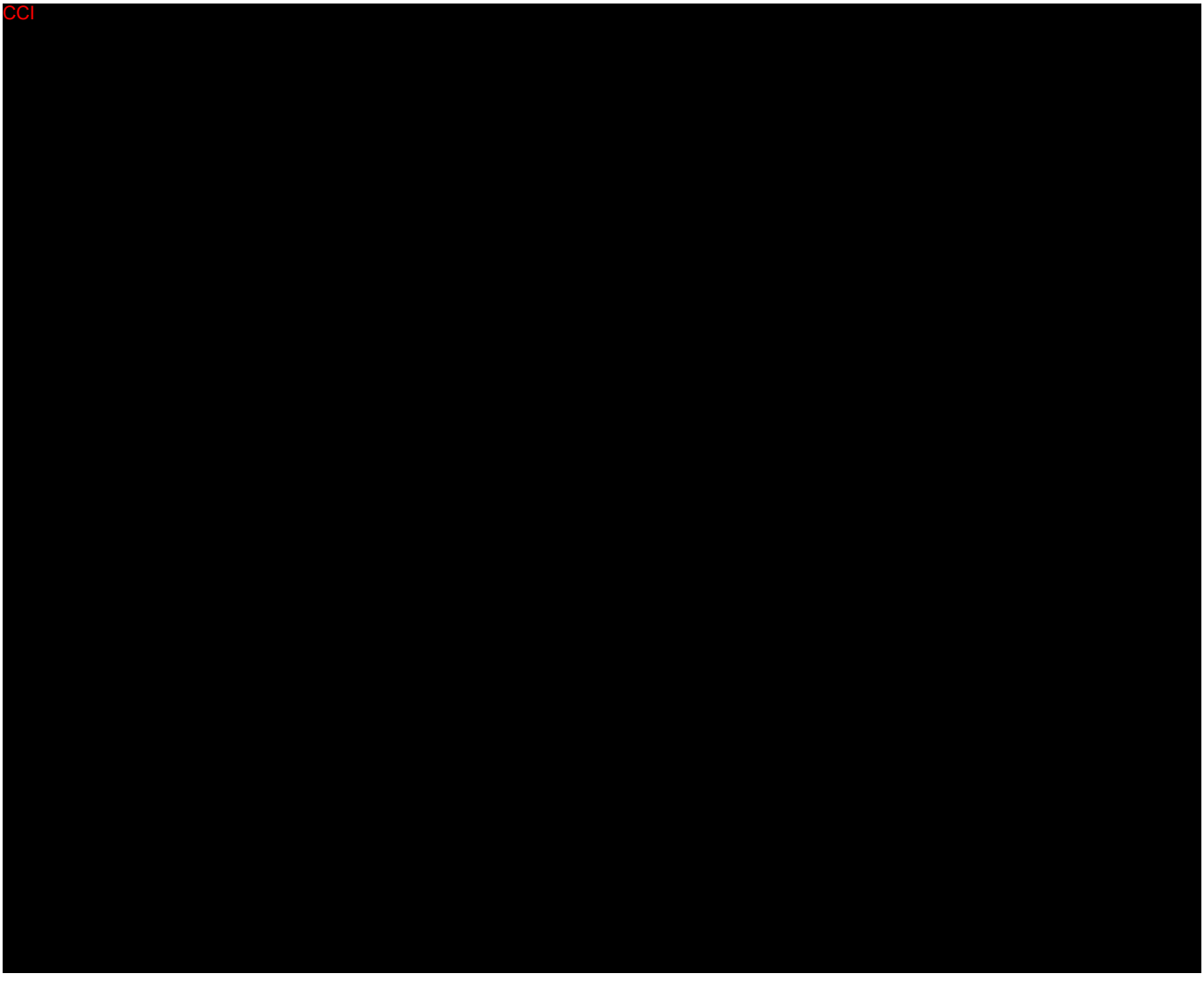
For analyses of the secondary endpoints, proportion-based efficacy variables will be using frequency tables as well as 95% two-sided Clopper-Pearson confidence intervals. With regard to time to event data, these will be summarized descriptively using Kaplan-Meier

methodology and plots, as well as median estimates based on Greenwood's formula, including 95% two-sided confidence interval. Kaplan-Meier analysis will be displayed in a table as well, giving the number of patients under risk, the number of events and the number of censoring for time intervals of 3 months duration. Kaplan-Meier Plots will present the number of patients under risk for each time point of the x-axis. Kaplan-Meier Plots and matching tables will be created for the subgroups

- Presence of liver metastases at baseline
- Primary tumor
- Primary side of tumor only.

Continuous endpoints will analyzed using summary statistics tables including 95% two-sided confidence intervals.

CCI



6.5 Safety

No formal statistical tests will be done for the safety endpoints. All analyses for safety will be performed in the FAS population.

6.5.1 Extent of exposure

Extent of exposure will be summarized for the FAS, using descriptive statistics.

Duration of study treatment will be calculated in days and presented in months as the date of the last dose of any study treatment – date of the first dose of any study treatment + 1.

Dose modification, if applicable, will be summarized for both treatments separately.

6.5.2 Adverse events

All adverse events (AE) whether considered drug-related or not, will be reported on the CRF with diagnosis, start/stop dates, dates of any grade change, action taken, whether treatment was discontinued, any corrective measures taken, and outcome. For all events, the relationship to treatment and the severity of the event will be determined by the Investigator. AEs will be classified and coded using the National Cancer Institute (NCI) Common Terminology Criteria Adverse Event (CTCAE), version 5.0 .

According to the protocol (See section 8.3.1 of the final study protocol) AE will be collected from the signing of the informed consent form (ICF) until the end of active FU visit, which should be 30 days after last dose of regorafenib and 100 days after last dose of nivolumab. An AE is considered as treatment-emergent (TEAE) if it starts during treatment or within the post-treatment time window.

For further definitions of the terms AE, SAE, seriousness, intensity, causal relationship with treatment, causal relationship to protocol-required procedures, action taken, and outcome please refer to the respective section of the protocol.

In the CRF only the worst action taken with regard to the study medication was captured, if nothing else changed while the AE was ongoing. Therefore all tables using this field, will have a footnote, mentioning this fact.

In addition to standard medDRA grouping, treatment-emergent AE of importance will be defined as follows:

- Rash grade 3 or higher. Rash consists of the following MedDRA LLT*:

  10037893, 10037894, 10005876, 10005877, 10012483, 10012884, 10013784, 10013785, 10058692, 10015138, 10015139, 10015143, 10015241, 10015242, 10015243, 10015244, 10015283, 10015285, 10015286, 10015585, 10015587, 10058694, 10058693, 10058721, 10015724, 10015725, 10016064, 10018081, 10018087, 10077181, 10018093, 10018094, 10018095, 10062432, 10062439, 10077163, 10062442, 10062443, 10049201, 10019117, 10024426, 10024437, 10079448, 10062704, 10079449, 10062705, 10025418, 10025422, 10049195, 10025423, 10025424, 10027022, 10027967, 10027968, 10027969, 10027970, 10028837, 10075807, 10074762, 10068748, 10033726, 10033727, 10033728, 10033729, 10033730, 10033737, 10033740, 10034120, 10054770, 10078918, 10074684, 10037844, 10037845, 10037872, 10048332, 10037848, 10037851, 10037855, 10048310, 10037856, 10037858, 10037863, 10037867, 10037868, 10037869, 10037870, 10037873, 10079349, 10037874, 10037875, 10037876, 10037878, 10037880, 10063646, 10066739, 10057984, 10037890, 10059872, 10038183, 10038186, 10038187, 10038188, 10081713, 10082785, 10080343, 10062880, 10040806, 10081571, 10040841, 10040913, 10081143, 10073477, 10047983, 10047984

- High Billirubin (if entered as AE)
- High Liver Enzymes (excluding GGT) (if entered as AE)
    - This table will be broken out by presence of liver metastases (yes vs. no) additionally
- Hand/Foot Syndrom, consisting of the following MedDRA LLT*:

  10033552, 10033552, 10070580, 10069640, 10019111, 10048776, 10019126, 10063069, 10033553, 10054524, 10070580, 10069640, 10019111, 10048776, 10019126, 10063069, 10033553, 10054524

All tables for TEAE/TESAE, except the one displaying (S)AE by relationship, will also be broken out by relationship to drug (Rego/Novo/Both). Relationship to drug in this case is defined as:

- Related to Rego: relationship to regorafenib = yes AND relationship to nivolumab = no
- Related to Nivo: relationship to nivolumab = yes AND relationship to regorafenib = no
- Related to Both: relationship to regorafenib = yes AND relationship to nivolumab = yes
- Any related: relationship to regorafenib = yes OR relationship to nivolumab = yes

Descriptive summary tables (frequency and percentage of patients, not of events) will be presented using the most recent version of MedDRA terminology, for the following:

- Overview over treatment-emergent AEs
- Treatment-emergent AEs with grade 3, 4, or 5
- Treatment-emergent AEs by worst CTCAE grade
- Treatment-emergent AEs occurring in at least 5% of patients
- Overview over treatment-emergent AEs leading to study drug withdrawal
- Treatment-emergent AEs leading to study drug withdrawal by worst CTCAE grade
- Treatment-emergent AEs leading to regorafenib dose reduction
- Treatment-emergent AEs leading to drug interruption
- Treatment-emergent AEs leading to dose reduction and/or drug interruption
- Treatment-emergent AEs of importance
- Overview over treatment-emergent drug-related AEs
- Treatment-emergent drug-related AEs with grades 3, 4, or 5
- Treatment-emergent drug-related AEs by worst CTCAE grade
- Treatment-emergent drug-related AEs occurring in at least 5% of patients
- Overview over treatment-emergent drug-related AEs leading to study drug withdrawal

- Treatment-emergent drug-related AEs leading to study drug withdrawal by worst CTCAE grade
- Treatment-emergent drug-related AEs leading to regorafenib dose reduction
- Treatment-emergent drug-related AEs leading to drug interruption
- Treatment-emergent drug-related AEs leading to dose reduction and/or drug interruption

The incidence of pre-treatment AE and non-treatment-emergent AEs during long-term follow-up will be tabulated. Other non-treatment-emergent AE will be presented by means of a subject data listing only.

Overview tables of adverse events will be presented in addition to the frequency tables.

6.5.3 Deaths and Serious Adverse events

Serious adverse events (SAE) will be classified using Version 5 of the National Cancer Institute (NCI) Common Terminology Criteria Adverse Event (CTCAE) grading system, and MedDRA terminology.

- Overview over treatment-emergent SAEs
- Treatment-emergent SAEs by worst CTCAE grade
- Treatment-emergent SAEs leading to study drug withdrawal
- Treatment-emergent SAEs leading to study termination
- Treatment-emergent SAEs leading to regorafenib dose reduction
- Treatment-emergent SAEs leading to drug interruption
- Treatment-emergent SAEs leading to dose reduction and/or drug interruption
- Overview over treatment-emergent drug-related SAEs
- Treatment-emergent drug-related SAEs by worst CTCAE grade
- Listing of treatment-emergent SAEs
- Listing of non-treatment-emergent SAEs

The incidence of deaths in the study and especially deaths up to within 30 and 100 days of last dose of Regorafenib and Nivolumab respectively will be summarized by cause of death. All deaths up to within 30 (100) days of last dose of study drug will be listed by patient with start and stop date of study medication, date of death, and cause of death. All deaths beyond 30 (100) days after last dose of study drug will be displayed in a separate listing.

6.5.4 Pregnancies

All pregnancies reported during the study will be listed.

6.5.5 Clinical laboratory data

Descriptive statistics (number of observations, mean, standard deviation, minimum, median and maximum values) will be presented for clinical laboratory tests (hematology, clinical biochemistry and urinalysis), their changes from baseline (including baseline value), and their percent changes from baseline at applicable visits.

Severity of hematological and biochemical laboratory values will be assessed based on NCI CTCAE Version 5.0. CTCAE severity grading for laboratory abnormalities are based on applicable laboratory threshold values outlined in NCI CTCAE. It should be noted that in the present analysis of those laboratory parameters for which additional clinical information potentially can also influence the toxicity grade, this clinical information is in general not available and only the laboratory measurements are used for severity assessment.

Any additional specific handling of the NCI CTCAE toxicity grading assignments will be noted in the footnotes of the corresponding tables as applicable per the data collection in the study.

- In the event of overlapping CTCAE criteria ranges for specific lab tests, the algorithm assigns the worst grade
- If calcium type is not recorded (i.e. only "calcium" is recorded), then grading is done as if the calcium is total calcium. "Calcium corrected" is computed from total calcium and serum albumin (if ≤ 4.0 g/dl) from the same time point based on CTCAE v5.0 guidance. If serum albumin (if ≤ 4.0 g/dl) from the same time point is not available or if "calcium, unspecified" was collected then grading is done as if the calcium is "corrected calcium."
- Results with special characters (such as ">" and "<") are not graded.

The frequency of treatment emergent laboratory abnormalities regarding hematology, coagulation panel, clinical chemistry, and urinalysis will be tabulated. Worst grades for hematological and biochemical toxicities will be calculated according to CTCAE based on laboratory measurements, and will be summarized by NCI CTCAE category and worst grade.

Clinical laboratory toxicities during treatment including a period of 30 days after last dose of treatment will be considered as "treatment-emergent" for Regorafenib. The duration for Nivolumab will be 100 days.

The last non-missing value before or on the first day of study drug will be retained as "baseline" data. If several assessments are performed on the same day (without timing information) the average of the values will be considered.

Incidence tables (frequency and percentage of patients) as well as tables with change in NCI CTCAE worst grade from baseline will be presented as following:

- Hematological and biochemical toxicity during screening (the last evaluation available before assignment to treatment is taken into account).
- Treatment-emergent hematological and biochemical toxicity.
- Treatment-emergent hematological and biochemical toxicities with incidence rate above 5% in any treatment arm.
- Change in worst grade for hematological and biochemical toxicity from baseline.

The laboratory values will be also categorized into low, normal and high according to their reference ranges.

Descriptive statistics will be calculated by time interval.

Unscheduled laboratory data will be listed (Section 16 of the CSR) only.

6.5.5.1 Laboratory abnormalities

Incidence tables for the following laboratory abnormalities under Nivolumab treatment will be summarized:

- ALT or AST > 3 x ULN
- ALT or AST > 5 x ULN
- ALT or AST > 10 x ULN
- ALR or AST > 20 x ULN
- Total bilirubin > 2 x ULN
- ALP > 1.5 x ULN
- Concurrent (within 1 day after Nivolumab treatment) ALT or AST > 3 x ULN and total bilirubin > 2 x ULN
- Concurrent (within 30 days after Nivolumab treatment) ALT or AST > 3 x ULN and total bilirubin > 2 x ULN
- TSH value > ULN and
    - with baseline TSH value <= ULN
    - with at least one FT3/FT4 test value < LLN within 2-week window after the abnormal TSH test
    - with all FT3/FT4 test values >= LLN within 2-week window after the abnormal TSH test
    - with FT3/FT4 missing within 2-week window after the abnormal TSH test
- TSH < LLN and
    - with baseline TSH value >= LLN
    - with at least one FT3/FT4 test value > ULN within 2-week window after the abnormal TSH test
    - with all FT3/FT4 test values <= ULN within 2-week window after the abnormal TSH test
    - with FT3/FT4 missing within 2-week window after the abnormal TSH test

6.5.6 12-Lead ECG, QTc

Analyses of ECG and QTc will be performed on the SAF population.

Corrected QT (QTc) will be calculated using Bazett's (QTcB) and Fridericia's (QTcF) formula. If several assessments are performed on the same day (without timing information) the average of the values will be considered.

Descriptive statistics including arithmetic mean, SD, median, minimum, and maximum will be presented for the following ECG parameters: HR, RR, PR, QT, QRS, QTcB and QTcF. Parameters will be summarized for actual results and the change from baseline for the safety analysis set at each scheduled visit.

The number and percent of patients with absolute QTc interval prolongation or QTc increase from baseline will be defined and summarized. The criteria are:

- QTc prolongation: QTc interval > 450 msec, QTc interval > 480 msec and QTc interval > 500 msec.
- Increase from baseline in QTc interval: QTc interval increases from baseline > 30 msec and QTc interval increases from baseline > 60 msec.

The number and percent of patients with new clinically significant abnormalities on ECGs per the investigator's assessment at post-baseline time points will be summarized.

Unscheduled ECG data will be listed (Section 16 of the CSR) only.

6.5.7 ECOG status

The ECOG performance status will be summarized by descriptive statistics tables over time, incl. change from baseline. Two sets of tables will be produced, interpreting the score as categorical variable (categories 0-5) as well as continuous variable.

6.5.8 Other safety measures

Vital signs (i.e. blood pressure, heart rate, weight and BMI) will be tabulated and summarized by visit for observed values and changes from baseline using descriptive statistics, as appropriate.

Unscheduled vital signs data will be listed (Section 16 of the CSR) only.

7. Document history and changes in the planned statistical analysis

This is the first amendment of the SAP.