| | |
|---|---|
| **Document Type:** | Statistical Analysis Plan |
| **Official Title:** | A randomized, open label, multicenter Phase 2/3 study to evaluate the efficacy and safety of rogaratinib (BAY1163877) compared to chemotherapy in patients with FGFR-positive locally advanced or metastatic urothelial carcinoma who have received prior platinum-containing chemotherapy |
| **NCT Number:** | NCT03410693 |
| **Document Date:** | 19 AUG 2020 |

# A randomized, open label, multicenter Phase 2/3 study to evaluate the efficacy and safety of rogaratinib (BAY 1163877) compared to chemotherapy in patients with FGFR-positive locally advanced or metastatic urothelial carcinoma who have received prior platinum-containing chemotherapy

**[Phase 2/3 study of rogaratinib (BAY 1163877) vs chemotherapy in patients with FGFR-positive locally advanced or metastatic urothelial carcinoma]**

| | | | |
|---|---|---|---|
| Bayer study drug | BAY 1163877 / rogaratinib | | |
| [Study purpose:] | Assess the efficacy and safety of rogaratinib compared to chemotherapy | | |
| Clinical study phase: | Phase 2/3 | Date: | 17MAY2019 |
| Study No.: | 17403 | Version: | Final 1.0 |
| Author: | PPD | | |



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

## Abbreviations

| | |
|---|---|
| AE | Adverse event |
| AESI | AE of special interest |
| AJCC | American Joint Committee on Cancer |
| ALT | Alanine aminotransferase |
| ANCOVA | Analysis of Covariance |
| AP | Asia Pacific |
| AST | Aspartate aminotransferase |
| ATC | Anatomical Therapeutic Chemical |
| AUC | Area under curve |
| b.i.d. | Twice daily, *bis in die* |
| BMI | Body mass index |
| CI | Confidence interval |
| CMH | Cochran-Mantel- Haenszel |
| CR | Complete Response |
| CRF | Case Report Form |
| CSR | Clinical Study Report |
| CT | Computed tomography |
| CTCAE | Common Terminology Criteria for Adverse Events |
| DCR | Disease-control rate |
| dL | Deciliter |
| DOR | Duration of response |
| DMC | Data Monitoring Committee |
| EORTC | European Organisation for Research and Treatment of Cancer |
| e.g. | For example, *exempli gratia* |
| ECOG | Eastern Cooperative Oncology Group |
| EOT | End of treatment |
| EQ-5D-3L | EuroQol Group's five dimensions questionnaire |
| FGFR | Fibroblast growth factor receptor |
| FAS | Full Analysis Set |
| GCP | Good Clinical Practice |
| GMS | Global Medical Standards |
| GFR | glomerular filtration rate |
| HEOR | Health Economics, Outcomes & Reimbursement |
| HRQoL | Health-related Quality of Life |
| HLT | Higher level term |
| ICH | International Committee on Harmonization |
| ICF | patient information/informed consent form |
| ITT | Intent to treat |
| i.v. | Intravenous(ly) |
| IxRS | Interactive voice/web response system |
| i.e. | That is, *id est* |
| kg | kilograms |
| KM | Kaplan-Meier |
| LKAD | Last known alive date |
| LOS | Listing only set |
| MedDRA | Medical Dictionary for Regulatory Activities |
| m | Meter |
| mg | Milligram |
| MRI | Magnetic resonance imaging |
| M&S | Modelling and Simulation |
| N/A | Not applicable |
| NCI | National Cancer Institute |

| | |
|---|---|
| NE | Non evaluable |
| OEE | Overall extent of exposure |
| ORR | Objective response rate |
| OS | Overall Survival |
| PI | patient information |
| p.o. | Orally, *per os* |
| PFS | Progression-free survival |
| PIK3CA | Phosphatidylinositol-3-kinase catalytic subunit alpha |
| PR | Partial Response |
| PRO | Patient-Reported outcome |
| PT | Preferred Term |
| RAS | Rat sarcoma |
| RECIST v.1.1 | Response evaluation criteria in solid tumors Version 1.1 |
| RNA-ISH | RNA *in situ* hybridization |
| QLQ-C30 | 30 item Quality of Life Questionnaire |
| QoL | Quality of Life |
| SAE | Serious adverse event |
| SAF | Safety analysis set |
| SAP | Statistical Analysis Plan |
| SAS | Statistical Analysis Software |
| SOC | System Organ Class |
| SD | Stable disease |
| TA | Therapeutic Area |
| TEAE | Treatment-emergent adverse event |
| TNM | Classification of Malignant Tumors (T = tumor, N = lymph node, M = metastasis) |
| ULN | Upper limit of normal |
| WHO-DD | World Health Organization Drug Dictionary |
| WT | Wild type |

# 1. Introduction

Rogaratinib (BAY 1163877) is an oral pan-FGFR inhibitor. It shows potent FGFR1, -2, -3 and -4 inhibition in biochemical assays that translates into strong inhibition of cellular downstream pERK resulting in inhibition of FGF2-stimulated tumor cell proliferation. Inhibition of cell proliferation by rogaratinib strongly correlates with expression of mRNA of FGFR isoforms as well as activation by somatic mutations. The high in-vitro potency and mode of action translates into strong in-vivo efficacy in tumor models which have an activated FGFR signaling in non-small-cell lung cancer, small-cell lung cancer, bladder, head and neck as well as breast cancer. In-vivo efficacy is correlated with effective inhibition of tumor FGFR phosphorylation and inhibition of FGFR downstream signaling as shown by pERK inhibition. Moreover, rogaratinib exhibits strong in-vivo anti-tumor efficacy in monotherapy in FGFR mRNA overexpressing xenograft models with good tolerability. Efficacy in inhibition of tumor growth as well as the respective mode of action of inhibition of FGFR signalling such as phosphorylation of ERK1/2 was also strongly correlated with abundant expression levels of FGFR isoforms in tumor cells or tumor tissues. In addition, rogaratinib demonstrated additive activity with standard-of-care therapy in lung and urothelial carcinoma models.

So far, clinical experience with rogaratinib is available from two clinical trials, study 16443 and study 16958. The encouraging results of Study 16443 support the continuation of the development

of rogaratinib in urothelial carcinoma. Study 17403 forms part of the sponsor' s clinical development program for rogaratinib. It is intended to generate the pivotal data in support of rogaratinib' s registration for the urothelial carcinoma indication in FGFR positive patients.

This Statistical Analysis Plan (SAP) v1.0 is based on the following documentation:

| Document | Date | Version |
|---|---|---|
| Protocol | 29November 2017 | 1.0 |
| Protocol Amendment (global amendment) | 09 May 2019 | 2.0 |

## 2. Study Objectives

The objectives of the study are provided below:

Primary objective:

- To demonstrate the superiority of rogaratinib over chemotherapy in terms of prolonging overall survival of urothelial carcinoma patients with FGFR positive tumors.

Secondary objectives:

- To evaluate additional efficacy including the following variables
    - progression-free survival (PFS)
    - objective response rate (ORR)
    - disease control rate (DCR)
    - duration of response (DOR)
- To evaluate the safety of rogaratinib (adverse events)

Tertiary objectives:

- Patient-reported outcome (PRO)
- Evaluate biomarkers to investigate the drug (i.e. mode-of-action-related effect and/or safety) and/or the pathomechanism of the disease
- Pharmacokinetics

The objective for the **Phase 2** part of the study is to demonstrate the early signal of efficacy of rogaratinib over chemotherapy in terms of objective response rate of urothelial carcinoma patients with FGFR positive tumors in PIK3CA/RAS WT (referred as WT hereafter) patients.

## 3. Study Design

This is a randomized, open-label, multicenter Phase 2/3 study to evaluate the efficacy and safety of rogaratinib (BAY 1163877) compared to chemotherapy in patients with FGFR-positive locally advanced or metastatic urothelial carcinoma who have received prior platinum-containing chemotherapy.

At randomization, patients will have locally advanced or metastatic urothelial carcinoma and have received at least one prior platinum-containing chemotherapy regimen. Only patients with FGFR1 or 3 positive tumors can be randomized into the study. Archival tumor tissue is adequate for testing of FGFR1 and 3 mRNA expression, which will be determined centrally using an RNA *in situ* hybridization (RNA-ISH) test.

All tumor images need to be submitted continuously for independent central imaging review.

## 3.1 Study periods and duration

The study will comprise the following periods:

1) Pre-treatment period, including FGFR testing and screening;

2) Treatment period;

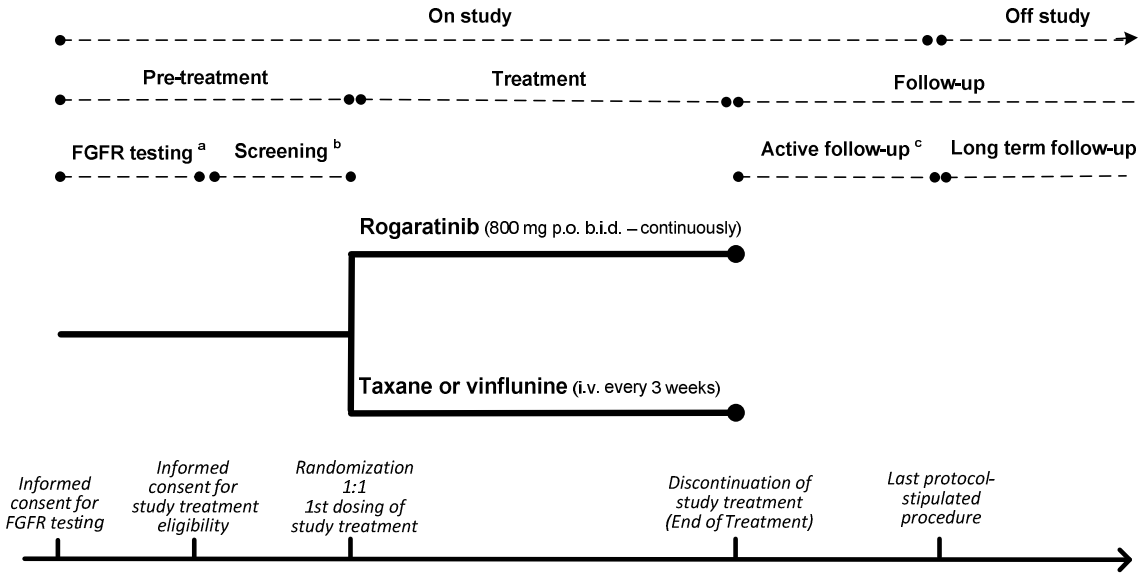
3) Active Follow-up period: including a safety follow-up at least 30 days ( up to +7 days) after the last administration of study treatment), and both safety and efficacy information will be collected for patients who discontinue study treatment without diseases progression;

4) Long-term follow-up period: patients will be contacted every month(+/- 7 days) to collect survival status and subsequent systemic anti-cancer treatment.

The Phase 2 and 3 parts of the study follow the same screening, treatment and follow-up assessment schedule.

Patients will be considered “on study” during the pre-treatment, treatment and active follow-up periods. During the long-term follow-up period, the patients will be considered “off study” (i.e. no study-related procedures with the patient). An overview of the study design is presented in Figure 3-1.

**Figure 3-1: Study 17403 - Design overview**




b.i.d. = Twice daily, *bis in die*; FGFR = Fibroblast growth factor receptor; i.v. = intravenously; mg = Milligram; PIK3CA = Phosphoinositide 3-kinase, catalytic subunit alpha isoform; p.o. = Orally, *per os*; RAS = Rat sarcoma; mRNA = Messenger ribonucleic acid.
a: During FGFR testing period, patients will be tested for FGFR1 and 3 mRNA expression levels and, after a positive FGFR test result was obtained, these patients are also tested for presence or absence of potential FGFR inhibitor resistance mutations, PIK3CA and/or RAS. PIK3CA and/or RAS mutation status will not affect patient eligibility, but will be used for patient stratification.
b: Only FGFR-positive patients (high expression of at least FGFR1 or 3) can enter screening.
c: Safety information is collected for all discontinued patients for at least 30 (up to +7) days after the last administration of study treatment, and both safety and efficacy information is collected for patients who discontinue study treatment without disease progression

### 3.2 Schedule of procedures

Refer to Protocol Table 9-1 for the detailed schedule for the assessments.

### 3.3 Determination of sample size

Refer to Protocol Section 10. 4 for the details about sample size calculation.

## 4. General Statistical Considerations

### 4.1 General Principles

Statistical analyses will be conducted by or under the supervision of the sponsor's Study Statistician, except for the analysis of biomarker data and pharmacokinetics/pharmacodynamics data, which will be performed by or under the direction of the sponsor's Genomics and Biomarker Statistical Expert, and PK experts. The details of those analysis will be described in a separate analysis Plan. and results will be provided in a separate report outside of the clinical study report (CSR).

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 continuous data. Frequency tables will be generated for categorical data.

All analyses for population characteristics (Section 6.1) and safety (Section 6.3) will be performed for WT patients first and will also be repeated for all study population if the test for primary endpoint OS in WT population is positive. For the order of efficacy analyses, refer to Section 6.2 for the details.

## 4.2 Handling of Dropouts

A patient who discontinues study participation prematurely for any reason is defined as a "dropout" if the patient has already been randomized, even if no study drug has been taken, patients who drop out will not be replaced.

Patients must be withdrawn from the study treatment, active follow-up or long-term follow-up if any of the withdrawn criteria is met. See Protocol Section 6.4.1.3 for pre-specified withdrawn criteria.

A subject who terminates the study before randomization but after completing Pre-screening (FGFR testing) is regarded as a 'screening failure'. A subject who fails to complete pre-screening (FGFR testing) period is a 'pre-screening failure'. All patients will enter the active follow-up period upon discontinuation of the study treatment, except for those who object to follow-up data collection. Safety information is collected during the Active Follow-up for all discontinued patients for at least 30 (up to +7) days after the last administration of study treatment. In addition, for those patients who discontinue study drug due to reasons other than death or PD (In case disease progression is not confirmed by central review, it is strongly recommended that patients go into active follow-up for tumor assessments until central imaging has confirmed disease progression), active follow-up assessments including both efficacy and safety information shall continue until disease progression is documented or new anti-tumor treatment is administered, whichever occurs first (and obviously death). Tumor assessments from the Active follow-up will also be used for primary and secondary analyses.

## 4.3 Handling of Missing Data

In order to achieve the goal of a well conducted clinical trial according to Good Clinical Practice (GCP), every effort will be made to resolve incomplete or missing dates during the course of the study (i.e. edit checks, data cleaning / monitoring etc.). However, despite best efforts, it may be inevitable that missing or incomplete data are reported. All missing or partial data will be presented in the patient data listing as they are recorded on the Case Report Form (CRF). Unless specifically specified, missing data will not be carried forward or otherwise imputed in any statistical analysis.

The following rules will be implemented where appropriate so as not to exclude patients from statistical analyses due to missing or incomplete data:

### 4.3.1 Time-to-event variables

For time-to-event analyses, the censoring mechanism is assumed to be non-informative. Patients will be handled as right-censored in time-to-event analyses, if applicable.

Missing or unevaluable tumor assessments (including a scheduled assessment that was not done and an incomplete assessment that does not result in an unambiguous tumor response according to RECIST v.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 assessment and tumor response.

Overall survival (OS) of patients alive at the time of analysis will be censored at the last date they were known to be alive (last known alive date: LKAD).

The LKAD is derived from the main data sources. Data reported during cleaning for the time after the data cut will not be considered in the derivation of LKAD. The last available date across all key data panels will be picked as the LKAD, e.g. visit dates, exposure information, laboratory measurements, tumor assessment dates and disposition events or follow up assessments will be used to determine survival status.

If a death date is partially or completely missing, it will be imputed following the imputation rules below:

1. If there is an adverse event (AE) with outcome 'Death', the date of death will be imputed by the end date of the AE.
2. If there is no AE with outcome 'Death' and only the day of death is missing, it will be imputed by the first day of the month unless the LKAD is later than this date; in which case the LKAD will be used for imputation. If both day and month are missing, then the LKAD will be used for imputation.

If a tumor assessment date is partially missing, it will be imputed following the imputation rules below:

1. If only the day is missing, it will be imputed by the first day of the month.
2. If day and month are missing, it will be imputed by the date of the last previous tumor assessment plus 21 days unless the LKAD is earlier than this date; in which case the LKAD will be used for imputation.

### 4.3.2 Response rates

If a patient has no post-baseline tumor assessment available, i.e. the overall best response assessment is missing, the patient will be included into denominator for the calculation of objective tumor response rate (ORR) and disease control rate (DCR).

### 4.3.3 Adverse Event Start Date

The general principle for imputing incomplete or missing adverse event start dates will be to assume the start date occurs at the earliest time on-treatment, whenever possible.

If the start date is completely missing, then the start date will be imputed using the date of the first dose of study medication.

If the incomplete start date has day and month missing, then the following will be applied:

- If the year is same as the year of first dose date, then the day and month of the first dose date will be assigned to the missing fields.

- If the year is prior to the year of first dose date, then December 31 will be assigned to the missing fields.
- If the year is after the year of first dose date, then January 01 will be assigned to the missing fields.

If the incomplete start date has missing day only, then the following will be applied:

- If the month and year are same as the year and month of first dose date, then the day of first dose will be assigned to the missing day.
- If the month and year are before the year and month of first dose date, then the last day of the month will be assigned to the missing day.
- If the month and year are after the year and month of first dose date, then the first day of the month will be assigned to the missing day.

## 4.4 Interim Analyses and Data Monitoring

A Data Monitoring Committee (DMC) will be instituted in order to ensure ongoing safety of study patients with respect to a risk/benefit assessment during periodic data review meetings, review results from the planned interim analysis and provide a formal recommendation for continuation/termination of the study and monitor study conduct to ensure the overall integrity of trial in maintained.

Two interim analyses and one final analysis are planned for this study as described in Section 10.5 in the protocol and DMC charter.

Recommendation for trial continuation will be guided by the monitoring boundaries at the formal interim analysis as well as safety evaluations from all data review meetings according to the DMC charter. All analyses to be presented to DMC to aid recommendations for trial continuation will be explicated in the DMC charter.

The DMC will be explicitly asked to give recommendations on the continuation or termination of the trial. Decisions on trial termination, amendment or cessation of patient recruitment based on risk benefit assessments will be made after recommendations from the DMC have been assessed by the sponsor.

## 4.5 Data Rules

### 4.5.1 Time intervals

If time intervals are to be displayed other than days in statistical evaluations, then one year is considered to have 365.25 days (average length of a year, including leap years), one month is considered to have 30.4375 days (average length of a month, including leap years), one week is considered to have 7 days, and one cycle is considered to have 21 days (i.e. 3 weeks).

### 4.5.2 Baseline

**Randomization baseline value for efficacy parameters**: is defined as the last non-missing value on or before the date of randomization.

**Baseline value for safety parameters**: is defined as the last non-missing value on or before the date of first dose.

### 4.5.3 Repeated measures

If there are repeated measurements per time point (e.g. laboratory values, vital signs, etc.), the following rules will be used (unless otherwise specified):

- In case of repeated measurements at any post baseline time point, the first measurement at scheduled visits will be used. Unscheduled visits will be used, if there are no measurements at scheduled visits. If the latter is the case, the first unscheduled visit will be used.

### 4.5.4 Stratification

Randomization will be stratified by:

- PIK3CA and/or RAS activating mutations (presence vs. absence)
- Prior immunotherapy (yes vs. no)
- Modified 4-factor Bellmunt risk score (high vs. low)

The following rule will be applied for WT and mutant subgroups to potentially remove some stratification factors from the subgroups to avoid over-stratification: if the smallest group within one stratification factor (other than activating mutation status) is less than 15% of all randomized patients in either subgroup (450 patients for WT, and the corresponding number of mutant patients for mutant subgroup), and all study population (with the exception of always including activating mutation status for all population ), then this stratification factor will not be introduced into the model for the stratified statistical analysis of efficacy endpoints. As the majority of patients are anticipated to be randomized by the time of the planned formal interim analysis, this rule will be applied to both 2nd interim analysis and final analysis of OS for consistency.

In case of discrepancies between stratification factors entered in the interactive voice/web response system (IxRS) and information entered in CRF, the information from IxRS will be used for primary analysis.

### 4.5.5 Region

For demographics overview and subgroup analyses, patients are grouped together into regions. The following regions are defined:

- Asia

- North America/ Western Europe/ Israel/ Australia
- Rest of the world

### 4.6 Validity Review

The results of the Blinded Data Review Meeting will be documented in the Blinded Data Review Report and may comprise decisions and details relevant for statistical evaluation. Any changes to the statistical analysis prompted by the results of the Blinded Data Review meeting will be documented in an amendment and, if applicable, in a supplement to this SAP.

## 5. Analysis Sets

### 5.1 Assignment of analysis sets

Final decisions regarding the assignment of patients to analysis sets will be made during the review of study data and documented in the final list of important deviations, validity findings and assignment to analysis set(s) (see section 4.6).

**Full analysis set (FAS)**

The FAS population is defined as all randomized patients. Patients will be analyzed as randomized, meaning even if a patient was randomized and received no drug or if randomized and received incorrect study drug at any time, these patients will still be analyzed for efficacy endpoints according to the treatment to which they were randomized based on IXRS.

The FAS analysis set as a primary population for efficacy analysis is identical to the Intent-to-treat (ITT) analysis set.

**Safety analysis set (SAF)**

The SAF population is defined as all patients who received at least one dose of study treatment (rogaratinib or chemotherapy). Patients will be included in the analyses according to the actual treatment they received..

The SAF population will be used for the safety analysis..

## 6. Statistical Methodology

### 6.1 Population characteristics

#### 6.1.1 Disposition

The number of patients pre-screened, screened, randomized, treated and entered active and long-term follow up period will be summarized by treatment group and overall. In addition, the number of subjects pre-screened, screened and included in each analysis population will be displayed overall and by region, country and investigator.

The number of FGFR testing failures and screening failures who terminate the study before randomization during pre-treatment period with the primary reason for discontinuation will be presented. The number of patients discontinuing the treatment, active and long-term follow-up

periods with the primary reason for discontinuation will be presented by treatment group and overall.

### 6.1.2 Protocol deviation

The number of patients with important protocol deviations will be presented by deviation coded term, treatment group and overall.

### 6.1.3 Demographics and other baseline characteristics

Demographic variables and baseline cancer characteristics will be summarized by treatment group and overall for the FAS and Safety analysis set. Quantitative data will be summarized by arithmetic mean, SD, median, minimum and maximum. Frequency tables will be provided for qualitative data.

In addition, the frequencies of stratification factors will be summarized per IXRS data. The discordance for each stratification factor between CRF and IXRS data will be presented in a shift table.

Demographic variables include age, gender, race, ethnicity, region (Asia/North America/ Western Europe/ Israel/ Australia; Rest of the world), body weight(kg), body height(cm), body mass index (BMI), systolic and diastolic blood pressure, heart rate and temperature. Age and BMI will each be analyzed as continuous variable and in addition categorized with the following categories:

- Age group (years): 18 - <65, ≥65 - <85, ≥85
- BMI group (kg/m²): <18.5, ≥18.5- <30, ≥30

The following baseline cancer characteristics will be analyzed:

- Cancer Category
- Location of primary cancer
- Time from initial diagnosis to the date of the randomization (months)
- Time since first progression/ relapse (months) to randomization
- Type of assessment of first progression/relapse
- Time from most recent progression/ relapse to the date of randomization (months))
- Type of assessment of most recent progression/relapse
- Histology of tumor
- Staging at initial diagnosis / Staging system at initial diagnosis (concatenated)
- AJCC Grading at initial diagnosis
- Residual primary tumor status at study entry
- If primary tumor was resected, status of surgical margin
- Stage at study entry/ Staging system at study entry (concatenated)
- Number of lesions (number of target and non-target lesions) at baseline –investigator

- Number of lesions (number of target and non-target lesions) at baseline – central reader
- PIK3CA and/or RAS activating mutations from CRF data (presence vs. absence)
- Modified 4-factor Bellmunt risk score (0, 1, 2, 3, 4) and prognostic factors tested within 7 days before randomization
    - ECOG performance status (0 vs. 1)
    - Hemoglobin level (0: ≥ 10 g/dL vs. 1: < 10 g/dL)
    - Liver metastases (0: absent vs. 1: present)
    - Time from last systemic anti-cancer therapy dose (0: ≥ 90 days vs. 1: < 90 days)
- Smoking history

### 6.1.4 Medical history

Medical history findings will be coded by Medical Dictionary for Regulatory Activities (MedDRA) version 20.1 or later. Medical history will be presented for each MedDRA Primary System Organ Class (SOC),) and Preferred Term (PT) for FAS population by treatment group and overall.

### 6.1.5 Prior and concomitant medication, Prior local Anti-cancer therapy, Systemic Anti-cancer Therapy during follow-up, Diagnostic and Therapeutic procedures, and Radiotherapy

All investigator-reported non-study medications taken before and/or during the study will be coded by Anatomical Therapeutic Chemical (ATC) classification system according to the World Health Organization Drug Dictionary (WHO-DD) version 2017SEP.

Non-study medications taken before and/or during the study will be categorized as prior medications, concomitant medications, and post treatment medications.
A non-study medication that started before the first administration of study treatment is considered as prior. Post treatment medications are defined as non-study medications taken on or after the last dose of study treatment.

Concomitant medications include:

- Non-study medications with a start or stop date after the date of the first dose of study treatment but before the date of last study treatment

- Non-study medications that started prior to the first dose of study treatment and are ongoing after the start of study treatment;

Detailed classifications of prior, concomitant and post medication are illustrated in Table 6–1.

If a start date is missing, the medication will be assumed to start prior to first dose of study drug. If the end date is unknown and 'ongoing' was not checked in CRF, the medication will be assumed to end at the last visit date, death date or withdrawal from study date, whichever is the latest.

Medications with missing start and stop date but flagged as being ongoing at end of study will be considered to have started prior to start of study medication and ended after stop of study medication.

All prior, concomitant and post-treatment medications will be listed, including verbatim descriptions and coded terms, and flags for prior/concomitant/post-treatment medications.

Prior, concomitant, and post-treatment medications will be summarized by ATC class, subclass and preferred term. For each patient, multiple records of the same concomitant medication will be counted once within a drug class and generic drug name.. Note that the same medication can appear multiple times in the table as it can have several ATC codes.

Additional Summaries will also be produced for prior/concomitant/post-treatment anti-cancer therapies (local therapy, systemic therapy and radiotherapy). The number of patients with local and systemic anti-cancer therapy will be presented by intent of procedure and number of regimens, and also by type and number of regimens for each treatment group, respectively.

The number of patients with radiotherapy will be presented by location, intent of procedure and number of regimens, and also by location, type and number of regimens, respectively.

The diagnostic and therapeutic procedure will be summarized using frequencies of patients undergoing each type of procedure by treatment group and overall.

All the summaries will be generated for the FAS population.

**Table 6–1 Medication Classification**

| | Prior to study drug | Study drug started | Treatment with study drug | Study drug stopped | Follow up | **Prior Medication?** | **Concomitant Medication?** | **Post-treatment Medication?** |
|---|---|---|---|---|---|---|---|---|
| C1 | ■ | ■ | | | | **Yes** | **No** | **No** |
| C2 | ■ | ■ | ■ | ■ | | **Yes** | **Yes** | **No** |
| C3 | | ■ | ■ | ■ | ■ | **No** | **Yes** | **Yes** |
| C4 | ■ | ■ | ■ | ■ | ■ | **Yes** | **Yes** | **Yes** |
| C5 | | | | ■ | ■ | **No** | **No** | **Yes** |
| C6 | | ■ | ■ | ■ | | **No** | **Yes** | **No** |
| C7 | | ■ | | | | **No** | **Yes** | **No** |
| C8 | | | | ■ | | **No** | **Yes** | **No** |

C1= medication started before date of first dose of study drug and ended on or before the date of first study drug
C2= medication started before date of first dose of study drug and ended before or on the date of last study drug
C3= medication started on or after date of first dose of study drug and ended/ongoing after date of last study drug
C4= medication started before date of first study drug and ended/ongoing after date of last study drug administration
C5= medication started on or after date of last dose of study drug and ended/ongoing after date of last dose of study drug
C6= medication started on or after date of first study drug and ended before or on the date of last study drug administration
C7= medication started and ended on date of first dose of study drug
C8= medication started and ended on date of last dose of study drug

### 6.1.6 Extent of Exposure

Following randomization, rogaratinib (BAY 1163877) is planned to be taken orally (p.o.) twice a day (b.i.d.), on every day during a 21-day treatment cycle.

Following randomization, the assigned chemotherapy (monotherapy) is planed to be given as intravenous (i.v.) infusion once every three weeks (on day 1 of a 21-day cycle). The cycle ends with the next infusion.

Descriptive statistical summaries will be provided by treatment group for the following variables:

- Duration of treatment: will be calculated in months as (the date of the last dose of study treatment – date of the first dose of study treatment + 1 )/30.4375
- Number of cycles
- Number of doses for rogaratinib arm and number of infusions for chemotherapy arm
- Total amount of dose administered (sum of dosage of all doses administered in mg for rogaratinib, and mg/m2 for infusions)
- Percent of planned dose received = Total amount dose administered (mg) / Planned dose(mg) × 100%.

  The planned dose will be calculated as: sum of all planned doses for rogartinib arm, and all doses calculated for chemotherapy arm between the date of first dose and last dose of study treatment
- Number and percent of patients with dose modifications (i.e., number (%) of patients with at least one dose modification, at least one dose interruption/delay, at least one dose reduction, or at least one dose re-escalation)
- Total number of dose interruptions/delays, reductions and re-escalations.
- The reason for dose interruptions/delays, reductions and re-escalations.
- Number of interruptions/delays, reductions and re-escalations per patient

All the summaries will be done in SAF by treatment group .

As a general rule, and in accordance with the Oncology Therapeutic Area Standard, leading '0mg' (prior to the first positive amount of drug) and trailing '0 mg' records (not followed by any positive amount of drug), will not be included in the calculation of any drug duration or amount. Interruption without resumption of study treatment is not counted as an interruption.

## 6.2 Efficacy

Efficacy analyses will be performed as randomized in the FAS.

### 6.2.1 Primary Efficacy Variable

Primary efficacy variable for this study is overall survival(OS). OS is defined as the time from randomization to death due to any cause. See Table 6– below for detailed censoring rules.

**Table 6–2 Overall Survival censoring rules**

| Situation | End Date | Censored | Reason for Censoring |
|---|---|---|---|
| Death on or prior to database cut off date | Date of Death | No | N/A |
| Not known to have died as of database cut off date | Last known alive date (LKAD) on or prior to database cutoff | Yes | No Death |
| Lost to follow-up prior to database cutoff | Last known alive date (LKAD) | Yes | Lost to follow up |
| If no assessment was done after randomization, then the patient will be censored at randomization date. | Randomization date | Yes | No assessment after randomization |

## Primary analysis of OS

For the analyses of overall survival, a stratified log-rank test controlling type I error at a level of 0.025 (one-sided) will be conducted for the WT population first (step 1). A full alpha of 0.025 (one-sided) will be passed on to OS in all study population and some selected secondary endpoints (step 2) if and only if the null hypothesis of OS for WT population is rejected. In case of failing to reject null hypothesis of OS for WT population,

step 2 will not be tested. The following hypothesis will be tested for the PIK3CA and RAS WT patients first and then all study population by using a stratified log-rank test controlling type I error at a level of 0.025 (one-sided):

- $H_{0,OS}$: hazard ratio (rogaratinib / chemotherapy) $\geq 1$

  *versus*

- $H_{1,OS}$: hazard ratio (rogaratinib / chemotherapy) $< 1$

Details on controlling family-wise type I error will be specified in section 6.2.3.

In addition to the stratified log-rank tes for the PIK3CA and RAS WT patients first and then all study population, the hazard ratio (rogaratinib / chemotherapy) for OS and its 95% confidence interval will be calculated using the Cox model, stratified by the same factors as log-rank test stated above. Kaplan-Meier curves will be generated, and median survival time together with the 25$^{th}$ and 75$^{th}$ percentiles and associated 95% CIs will be presented. The KM estimates at time points such as 3 months, 6 months etc. together with corresponding 95% confidence intervals as well as the differences of these estimates will also be calculated between the rogaratinib arm and the chemotherapy arm.

## Sensitivity analysis of OS

WT and all study population: A sensitivity analysis will be performed by only including confirmed FGFR positive patients for PIK3CA and RAS WT population and all study population, respectively. All subject enrolled to this study are expected to have FGFR positive testing results. Due to inter-evaluator variability in the scoring of FGFR test , a confirmation test may be done at discretion of

the sponsor to identify the FGFR positive subjects which will be referred as confirmed FGFR positive.

Mutant subgroup: Three sensitivity analyses will be performed for mutant subgroup:

- First analysis will include all subjects with both known or unknown PIK3CA and RAS mutation status, per IxRS randomization scheme.
- Second analysis will only include subjects with known PIK3CA and RAS mutant.
- Third analysis will be performed for confirmed FGFR positive subjects with both known or unknown PIK3CA and RAS mutation status

In case of discrepancies between stratification factors entered in the interactive voice/web response system (IxRS) and information entered in CRF, sensitivity analyses may be performed for OS analysis based on information entered in CRF for both WT and all study population, respectively.

### 6.2.2 Secondary Efficacy Analysis

#### 6.2.2.1 Secondary efficacy endpoints

**Progression-free survival (PFS)** is defined as the time (days) from randomization to date of first observed disease progression (radiological or clinical assessment or both ) or death due to any cause(if death occurs before progression is documented). Clinical progression refers to symptomatic deterioration as defined in the protocol Section 6.4.1.3.

The actual date that the tumor scan was performed will be used for this calculation. If a tumor assessment is performed over more than one day (e.g. scans for chest and abdomen done on different days for a same evaluation), the earliest date will be used for the calculation of PFS.

**Censoring Rules**

The detailed censoring rules for PFS are described in Table 6– below.

Table 6–3 Censoring for progression-free survival(PFS)

| Situation | End Date | Censored | Reason for Censoring |
|---|---|---|---|
| No baseline radiological tumor assessments | Date of Randomization | Yes | No baseline assessment. |
| No post-baseline radiological assessment, or symptomatic deterioration or death | Date of Randomization | Yes | No post-baseline radiological assessment, symptomatic deterioration, or death |
| No progression and no death (with a post baseline tumor assessment) at data cut-off | Date of last tumor assessment prior or on data cut-off date | Yes | No Progression or death |
| Subject lost to follow-up or discontinued from study without progression or death | Date of last tumor assessment | Yes | Lost to follow-up or discontinued study due to a reason other than PD or death |

| | | | |
|---|---|---|---|
| New non-protocol permitted systemic anti-cancer treatment started before progression or death | Date of last radiological assessment before starting new systemic anti-cancer treatment | Yes | New systemic anti-cancer treatment started |
| Subjects discontinued from study due to PD, but no documented date of PD | Date of last radiological assessment | Yes | Subjects discontinued from study due to PD, but no documented date of PD |
| Death or progression after two or more consecutive missed radiological assessments* | Date of last adequate tumor assessment before missed assessments | Yes | Missed two or more consecutive tumor assessments |
| Subject had a progression after randomization (no 2 consecutive missing radiological assessments*) | Date of first progression event (symptomatic deterioration or radiological progression, whichever comes earlier). If a tumor assessment was done on multiple days, use the earliest date for that visit | No | N/A |
| Death without documented progression and without missing two or more consecutive radiological assessment prior to death | Date of Death | No | N/A |

* Two consecutive missing disease assessments is defined as the time interval in days between two consecutive tumor assessments. This number is greater than 140 days =2 × (63+7), where 63-day is the longest scheduled frequency for tumor assessment and 7-day is the protocol-allowed window.

**Note:** Progression in the table above refers to radiological progression or symptomatic deterioration

**Objective response rate (ORR)** is defined as the percentage of patients with complete response (CR) or partial response (PR). Patients for whom best overall response is not CR or PR, as well as patients without any post-baseline tumor assessment will be considered non-responders.

**Disease-control rate (DCR)** is defined as the percentage of patients, whose overall best response was not progressive disease (i.e. CR, PR, SD or Non CR/Non PD). Tumor assessments that is performed earlier than 6 weeks from the start of treatment will not be taken into account in the assessment of SD.

**Duration of response (DOR)** (for subjects with PR and CR only) is defined as the time from the first documented objective response of PR or CR, whichever is noted earlier, to disease progression (including symptomatic deterioration) or death whichever is earlier. The censoring rules for DOR are described in Table 6-4.

Table 6–4 Censoring for duration of response(DOR)

| Situation | End Date | Censored | Reason for Censoring |
|---|---|---|---|
| No progression and no death at data cut-off | Date of last tumor assessment before data cut-off | Yes | No Progression or death |
| Subject lost to follow-up or discontinued from study without | Date of last tumor assessment | Yes | Lost to follow-up or discontinued study due to |

| | | | |
|---|---|---|---|
| progression or death | | | a reason other than PD or death |
| New non-protocol permitted systemic anti-cancer treatment started before progression or death | Date of last radiological assessment before starting new systemic anti-cancer treatment | Yes | New systemic anti-cancer treatment started |
| Subjects discontinued from study due to PD, but no documented date of PD | Date of last radiological assessment | Yes | Subjects discontinued from study due to PD, but no documented date of PD |
| Death or progression after two or more consecutive missed radiological assessments* | Date of last adequate tumor assessment before missed assessments | Yes | Missed two or more consecutive tumor assessments |
| Subject had a progression after achieving CR or PR (no 2 consecutive missing radiological assessments*) | Date of first progression event (symptomatic deterioration or radiological progression, whichever comes earlier). If a tumor assessment was done on multiple days, use the earliest date for that visit | No | N/A |
| Death without documented progression and without missing two or more consecutive radiological assessment prior to death | Date of Death | No | N/A |

* Two consecutive missing disease assessments is defined as the time interval in days between two consecutive tumor assessments. This number is greater than 140 days =2 × (63+7), where 63-day is the longest scheduled frequency for tumor assessment and 7-day is the protocol-allowed window.

**Note:** Progression in the table above refers to radiological progression or symptomatic deterioration

#### 6.2.2.2 Analysis of secondary efficacy endpoints

Depending on study success in the primary efficacy variable in the WT subgroup, the secondary efficacy variables based on final database will be tested in the WT subgroup and all study population according to the multiple testing strategy outlined in Section 6.2.3.

The central reviewer will evaluate the radiological imaging produced by the investigator. The primary analysis of radiological endpoints (ORR, PFS, DCR, DOR) will be based on the central radiological reviewer assessments. Investigator assessments will be provided as sensitivity analysis.

**Analysis of PFS**

Rogaratinib and chemotherapy treatment arms will be compared using a log-rank test stratified by the same stratification factors as used in the analyses of the primary variable OS.

The hazard ratio (rogaratinib / chemotherapy) and 95% confidence intervals will be provided. KM estimates and KM curves will also be presented for each treatment arm. The KM estimates at time points such as 3 months, 6 months etc. together with corresponding 95% confidence intervals as well as the differences of these estimates between the rogaratinib arm and the chemotherapy arm will also be calculated.

**Analysis of ORR and DCR**

ORR will be compared between treatment arms using the Cochran-Mantel- Haenszel (CMH) test[1] adjusting for the same stratification factors as used in the analyses of the primary variable OS for both WT and all study population.

Estimates and 95% confidence intervals will be computed for each treatment arm. The differences in ORR between the rogaratinib and chemotherapy arms and the corresponding 95% confidence intervals will also be calculated.

Summary statistics will be displayed for all best response categories: CR, PR, SD, Non CR/Non PD, PD by central radiographic imaging, and PD by clinical judgment. Frequency counts and percentages with exact 95% confidence intervals will be displayed.

Same analyses stated above will be done for DCR based on independent central review as well.

**Analysis of DOR**

Since the responders are not a randomized group, no statistical testing will be performed for DOR. Analysis of DOR will be descriptive in nature based on central review assessments. KM estimates and KM curves will be displayed for each treatment arm.

### 6.2.2.3 Sensitivity Analysis of Secondary Efficacy Endpoints

The following sensitivity analysis will be performed:

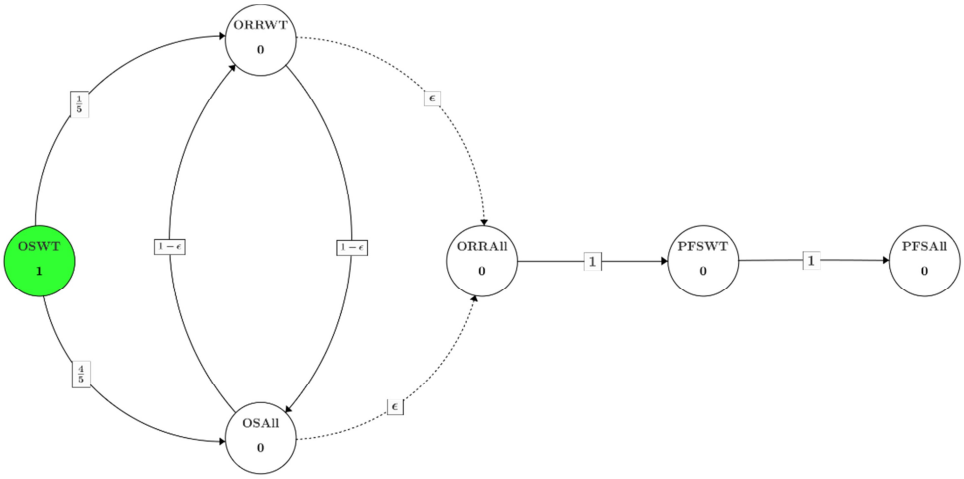
1) The analyses of ORR, PFS, DCR and DOR based on investigator's assessments will be conducted using the similar methodology as described in the section above.
2) The analyses of ORR, PFS, DCR and DOR for confirmed FGFR positive WT subjects, and all population
3) In case of discrepancies between stratification factors entered in the interactive voice/web response system (IxRS) and information entered in CRF, sensitivity analyses may be performed for all secondary analysis based on information entered in CRF.
4) For all secondary efficacy endpoints, two supportive analyses will be performed for PIK3CA and RAS mutant patients: First analysis will be performed by including mutant patients and those with unknown mutation status, per IxRS randomization scheme; Second analyses will only include known PIK3CA and RAS mutant patients.
5) For this study, in exceptional cases, when it is deemed clinically necessary by the investigator due to safety reasons, an imaging scan at an earlier time point than 6 weeks from

start of treatment for the 1st 'on treatment' tumor assessment is allowed. If extremely imbalanced such early scans occur between treatment arms, a sensitivity analysis will be performed for PFS to take the first assessment date as 6 weeks for both treatment groups, when they occur before 6 weeks.

### 6.2.3 Statistical test strategy for secondary endpoints

When targeted death events are reached in PIK3CA and RAS WT patients, final statistical analysis will be performed for both primary endpoint (OS) and key secondary endpoints (ORR, PFS) controlling family wise type I error rate (α) of one-sided 0.025.

A graphical summary of the testing strategy including primary and secondary endpoints mentioned above for PIK3CA and RAS WT patients and all study population is displayed in the figure below, using the methodology as outlined in "A graphical approach to sequentially rejective multiple test procedures" by Bretz F. et al[1].




In this testing strategy, OS in PIK3CA and RAS WT will be tested at significance level of one-sided 0.025 as the first step. If OS in PIK3CA and RAS wt patients is positive with the given significance level, OS in all study population (referred as "OS all" in the graph and below) will be tested (at significance level of one-sided 0.02, 4/5 of full α) in parallel with the ORR in PIK3CA and RAS WT patients (referred as ORR WT in the graph and below; at one-sided significance level of 0.005, 1/5 of full α). If either endpoint (OS all or ORR WT) is positive, the respective α level from the positive endpoint will be propagated to the other endpoint, which can be re-tested at full α level (except for a small close-to-zero ε, set to 0.0001 of full α). After test (and potentially re-test), if both OS all and ORR WT are positive, three endpoints, ORR in all study population (referred as ORR all in the graph), PFS in PIK3CA and RAS wt patients (referred as PFS WT in the graph) and PFS in all study population (referred as PFS all in the graph) will be tested at the significance level of one-sided 0.025 sequentially with the fixed sequence as just mentioned. In another words, among the three endpoints, only if the endpoint mentioned earlier in the sequence tested positive, can the latter endpoint be tested.

### 6.2.4 Subgroup Analyses

Subgroup analyses will be performed in a descriptive fashion for the primary efficacy endpoint (OS) as well as the secondary efficacy endpoints ORR, PFS, DCR and DOR. For time to event variables, hazard ratio estimates with 95% CI for PFS and DOR will be provided at least within each category of the subgroup, provided there are sufficient numbers of events within the subgroup across the treatment arms. Kaplan-Meier estimates of median times (including 95% CI) and Kaplan-Meier curves will be provided. For ORR and DCR, response rate estimates and 95% confidence intervals will be computed for each treatment arm within the subgroup.

Subgroup levels to be analyzed are:

- Planned dose for rogartinib arm (600 mg vs. 800 mg)
- Stratification factors at randomization (according to IxRS); However, in case of discrepancies between stratification factors entered in the interactive voice/web response system (IxRS) and information entered in CRF, the information from IxRS will be used for primary analysis but information from CRF will also be summarized
- Region (Asia vs. North America [NA]/Western Europe [W EU]/Israel/Australia vs. Rest of World [RoW])
- Age (<65 vs. ≥ 65, calculated at the date of randomization using date of birth)
- ECOG performance status (0 vs. 1)
- Gender (male vs. female)
- Race
- Ethnicity (Hispanic or Latino vs. non-Hispanic and non-Latino)
- Renal function (GFR <45 vs. 45 - <60 vs. ≥60 mL/min/1.73m²)
- hepatic function (AST and ALT each categorized by ≤ULN vs. >ULN ULN )

## 6.3 Safety

### 6.3.1 Adverse events (AEs)

Summary of adverse events is considered a secondary objective for this study. All AEs, whether considered drug-related or not, will be reported on the eCRF with 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. Adverse events will be coded using Medical Dictionary for Regulatory Activities (MedDRA) terms (version 20.0 or later), and graded according to National Cancer Institute's Common Terminology Criteria for Adverse Events (NCI CTCAE), version 4.03.

A treatment-emergent AE (TEAE) is defined as any event arising or worsening after start of study drug administration until 30 days after the last study drug intake (end of safety follow-up).

Descriptive summary tables (frequency and percentage of patients, not of events) will be presented by treatment arm and MedDRA terms and NCI CTCAE v.4.03 worst grade for the following:

- Incidence rate of pre-treatment AEs
- Incidence rate of all treatment-emergent AEs (TEAEs)
- Incidence rate of treatment-emergent drug-related AEs
- Incidence rate of treatment-emergent serious AEs (SAEs)
- Incidence rate of serious treatment-emergent drug-related AEs
- Interval-specific and cumulative event rates for the more common AEs for treatment emergent adverse events
- Incidence rate of treatment-emergent AEs of special interest (AESIs)

**Deaths and serious adverse events (SAEs)**

The incidence of deaths in the study, and especially deaths up to 30 days of last dose of study drug, will be summarized by treatment arm and cause of death. All deaths up to 30 days of last dose of study drug will be listed by patient with start and stop date of study treatment, date of death, and cause of death for both populations.

In addition to the above mentioned incidence table of all treatment-emergent SAEs, these events will be listed by patient with investigator AE term, start and stop date of study drug administration, start and stop date of AE, drug relationship, worst NCI CTCAE grade, action taken and outcome.

**AEs leading to discontinuation of study treatment and/or withdrawal from the study and other significant AEs**

The incidence of AEs leading to permanent discontinuation of study treatment and / or withdrawal from the study will be summarized by treatment arm for both populations, and be listed by patient with investigator AE term, start and stop date of study drug administration, start and stop date of AE, drug relationship, worst NCI CTCAE grade, action taken and outcome. In addition, incidence of AEs that caused dose reduction or interruption will be summarized separately by treatment arm.

**Adverse Events of Special Interest**

Signs and symptoms suggestive for soft-tissue mineralization, and retinal disorders (classified analog to CTCAE v.4.03 as Grade ≥ 2) are defined as AEs of special interest (AESI).

Soft-tissue mineralization related AESIs includes any events consistent with ectopic calcification (eg.MedDRA lowest level terms as following: tissue calcification, metastatic calcification, calcification of muscles, intestinal calcification, cutaneous calcification).

Any symptomatic retinal disorders including retinal detachment / retinal pigment epithelial detachment / serous retinopathy / retinal vein occlusion (classified analog to CTCAE v.4.03 as Grade ≥ 2) should be reported as AESI.

AESIs will be summarized by treatment arm, MedDRA terms and CTCAE v.4.03 worst grade. The corresponding listing will be provided as well.

### 6.3.2 Clinical laboratory evaluations

The following laboratory parameters will be summarized for the safety analysis set: hematology panel, coagulation panel, biochemistry panel, hormone panel, and urinalysis test.

Worst grades for hematological and biochemical toxicities will be calculated according to CTCAE v.4.03 based on laboratory measurements. It should be noted that in the process of assigning toxicity grades of those lab parameters for which additional clinical information potentially can also influence the toxicity grade, this clinical information is in general not available and only the lab measurements are used for toxicity grading

Hematological and biochemical laboratory toxicities assigned by the investigator that include clinical assessments are available in Adverse Events database and are summarized in Adverse Event tables.

Quantitative laboratory data will be described by the following summary statistics: number of observations, arithmetic mean, and standard deviation, median, minimum, and maximum. These summary statistics will be presented by treatment group for the original data as well as for the change from baseline.

Frequency tables will be provided for qualitative urinalysis data.

Laboratory data outside the reference range will be listed with abnormal values / alert ranges flagged. The incidence of laboratory data outside the reference range (low, high) will be summarized by treatment group in frequency tables.

For hematology and biochemical laboratory values, the incidence of laboratory toxicities will be summarized by worst CTCAE v4.03 grade and by treatment group. Shift tables will be also provided for the changes of worst CTCAE v4.03 grade after start of treatment *vs.* baseline.

### 6.3.3 Other safety measures

Quantitative data (vital signs and 12-Lead ECG) will be described by the following summary statistics: number of observations, arithmetic mean, standard deviation, median, minimum, and maximum. These summary statistics will be presented by treatment group for the original data as well as for the change from baseline if applicable.

Frequency tables will be provided for qualitative data (e.g. ECOG performance status, Pregnancy test , ECG interpretation, and overall assessment from ophthalmologic examination).

### 6.3.4 Pharmacokinetics/pharmacodynamics

The details of pharmacokinetic analysis and results will be described in the separate analysis plan and report.

### 6.3.5 Biomarker evaluation

Biomarker analyses and results will be provided in the separate analysis plan and report.

### 6.3.6 Patient-reported outcomes (PRO)

PRO data as measured by the EORTC QLQ-C30 and EQ-5D will be analyzed to assess differences in health-related Quality of Life (HRQoL) and health utility values between treatment arms based on time-adjusted Area Under the Curve (AUC) using all available data[2,3]. Summary statistics will be presented for each of the PRO endpoints at each assessment time point and for change from baseline by treatment group for both populations.

Primary Analysis of PROs: An analysis of covariance (ANCOVA) model will be used to compare the time-adjusted AUCs between the two treatment groups with covariates for baseline HRQoL score and the same stratification factors as used in the primary efficacy endpoint at the end of the study for both populations. Least-squares mean estimates; standard errors and 95% CIs will be estimated for each treatment group and for the treatment group difference for both populations. The treatment by covariate interactions will be explored in the ANCOVA models and the consistency of the treatment effect on HRQoL across different subsets defined by the covariates in the model will be assessed. In the event that the ANCOVA model assumptions (e.g., normality and homogenous variance of the error terms, equality of slopes for different treatment regression lines) are not satisfied, the rank analysis of covariance may be used [4, 5).

Secondary Analyses of PROs: Sensitivity analysis using different imputation methods for imputing missing assessments may be performed. Additional exploratory analyses may be carried out using the linear mixed effect models to explore the effects of treatment, time and other covariates on the endpoints, assuming the missing data mechanism is missing at random.

## 7. Document history and changes in the planned statistical analysis

None

## 8. References

1. Frank B., Willi M., Werner B, and Martin P. A graphical approach to sequentially rejective multiple test procedures. Statisics in Medicine, 28:586-604(2009).
2. EORTC. QLQ-C30 Scoring Manual. 2001 (Brussels);Third edition.
3. Oppe M RR, deCharro F. EuroQol EQ-5D User Guide. www.euroqol.org. 2007;Version 1.
4. Koch GG, Carr GJ, Amara IA, Stokes ME, Uryniak TJ. Categorical data analysis, in Statistical Methodology in the Pharmaceutical Sciences, ed. D.A. Berry, New York:Marcel Dekker Inc. 1990:391-475.
5. Quade D. Rank Analysis of Covariance. J Am Stat Assoc. 1967/12/01;62(320):1187-200

# Title page

## A randomized, open label, multicenter Phase 2/3 study to evaluate the efficacy and safety of rogaratinib (BAY 1163877) compared to chemotherapy in patients with FGFR-positive locally advanced or metastatic urothelial carcinoma who have received prior platinum-containing chemotherapy

**Phase 2/3 study of rogaratinib (BAY 1163877) vs chemotherapy in patients with FGFR-positive locally advanced or metastatic urothelial carcinoma**

| | | | |
|---|---|---|---|
| Bayer study drug | BAY 1163877 / rogaratinib | | |
| Study purpose: | Assess the efficacy and safety of rogaratinib compared to chemotherapy | | |
| Clinical study phase: | Phase 2/3 | Date: | 24 APR 2020 |
| Study No.: | 17403 | Version: | 1.0 |
| Author: | PPD | | |



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

## Abbreviations

| | |
|---|---|
| AE | Adverse event |
| AESI | AE of special interest |
| CRF | Case Report Form |
| CSR | Clinical Study Report |
| DCR | Disease-control rate |
| DOR | Duration of response |
| DMC | Data Monitoring Committee |
| ECOG | Eastern Cooperative Oncology Group |
| FGFR | Fibroblast growth factor receptor |
| IxRS | Interactive voice/web response system |
| PFS | Progression-free survival |
| PIK3CA | Phosphatidylinositol-3-kinase catalytic subunit alpha |
| PRO | Patient-Reported outcome |
| SAP | Statistical Analysis Plan |
| ULN | Upper limit of normal |
| WT | Wild type |

# 1. Introduction

This Statistical Analysis Plan (SAP) supplement is based on the operational changes reflected in Protocol Amendment 4 (Integrated Protocol v. 3.0) dated 12 NOV 2019. This Supplemental SAP version 1.0 is a supplement of main SAP version 1.0 dated 17 MAY 2019.

The study was originally designed as Phase 2/3 study with two interim analyses. However, before achieving the criteria for the first interim analysis at the end of Phase 2 part, the study recruitment was put on hold after a recommendation of the study DMC. The sponsor decided to perform a full analysis and review of all study data and stopped the recruitment permanently. Due to the stop of enrollment and full data analysis, the study will not move forward to its Phase 3 part and will remain as Phase 2.

This Supplemental SAP describes analyses that were not included in the main SAP but may be used for CSR. Changes to planned analyses in the SAP v1.0 are also described in this Supplemental SAP.

# 2. Study Objectives

Refer to main SAP v1.0 dated 17 MAY 2019 for the objectives of original study design.

The objective for the Phase 2 part of the study is to demonstrate the efficacy of rogaratinib over chemotherapy in terms of objective response rate of urothelial carcinoma patients with FGFR positive tumors. The Phase 3 part of the study will no longer be conducted, therefore OS will be considered as an exploratory efficacy variable for Phase 2.

# 3. Study Design

Refer to the Integrated Protocol v3.0 (dated 12 NOV 2019) Section 5 for study design.

## 3.1 Study periods and duration

Refer to main SAP v1.0 dated 17 MAY 2019.

## 3.2 Schedule of procedures

Refer to the Integrated Protocol v3.0 (dated 12 NOV 2019) Table 9-1 for the detailed schedule for the assessments.

## 3.3 Determination of sample size

Refer to the Integrated Protocol v3.0 Section 10.4 (dated 12 NOV 2019) for the details about sample size calculation.

# 4. General Statistical Considerations

## 4.1 General Principles

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.1.1 Changes in analyses for population characteristics and safety

Original planned analysis in main SAP v1.0:

All analyses for population characteristics (Section 6.1) and safety (Section 6.3) will be performed for WT patients first and will also be repeated for all study population if the test for primary endpoint OS in WT population is positive. For the order of efficacy analyses, refer to Section 6.2 for the details.

New proposed analysis:

Given that the study was terminated at Phase 2 part, OS is not the primary efficacy endpoint any more, therefore, previously proposed step-wise analysis of OS will not be performed. All analyses for population characteristics (Section 6.1) and safety (Section 6.3) will be performed for all study population. For demographic and baseline characteristics variables, summary will also be provided for WT population.

## 4.2 Handling of Dropouts

Refer to main SAP v1.0 dated 17 MAY 2019

## 4.3 Handling of Missing Data

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.3.1 Time-to-event variables

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.3.2 Response rates

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.3.3 Adverse Event Start Date

Refer to main SAP v1.0 dated 17 MAY 2019

## 4.4 Interim Analyses and Data Monitoring

Two interim analyses and one final analysis were originally planned for this study as described in Section 10.5 in the Integrated Protocol v3.0 (dated 12 NOV 2019) and DMC charter.

Before achieving the criteria for the first interim analysis, the study recruitment was put on hold after a recommendation of the study DMC. The sponsor decided to perform a full analysis and review of all study data and stopped the recruitment permanently.

## 4.5 Data Rules

### 4.5.1 Time intervals

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.5.2 Baseline

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.5.3 Repeated measures

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.5.4 Stratification

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.5.5 Region

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.6 Validity Review

Refer to main SAP v1.0 dated 17 MAY 2019

## 5. Analysis Sets

Refer to main SAP v1.0 dated 17 MAY 2019

## 6. Statistical Methodology

### 6.1 Population characteristics

### 6.1.1 Disposition

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.1.2 Protocol deviation

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.1.3 Demographics and other baseline characteristics

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.1.4 Medical history

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.1.5 Prior and concomitant medication, Prior local Anti-cancer therapy, Systemic Anti-Cancer Therapy during follow-up, Diagnostic and Therapeutic procedures, and Radiotherapy

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.1.6 Extent of Exposure

Changes for calculating duration of treatment for Chemotherapy arm:

Old algorithm:

Duration of treatment: for rogaratinib arm: will be calculated in months as (the date of the last dose of study treatment – date of the first dose of study treatment + 1)/30.4375

New algorithm:

Duration of treatment: for rogaratinib arm: will be calculated in months as (the date of the last dose of study treatment – date of the first dose of study treatment + 1)/30.4375. for chemotherapy arm: will be calculated in months as: (the date of the last dose of study treatment – date of the first dose of study treatment + 21)/30.4375

New analysis added:

The overall treatment duration will be summarized under 800mg, 600mg and 400 mg dosage for rogaratinib arm.

## 6.2 Efficacy

### 6.2.1 Changes in Primary Efficacy Variable

Overall survival (OS) was the primary efficacy variable for the planned Phase 3 part. Given that the study will not move forward to Phase 3 part, the primary efficacy variable for Phase 2 part is objective response rate (ORR) based on central assessment. Overall survival will be considered as exploratory efficacy variable for Phase 2.

**Primary analysis of ORR**

ORR will be compared between treatment arms using Fisher exact test for both WT and all study population.

**Sensitivity analysis of ORR**

1) The analyses of ORR based on investigator's assessments will be conducted for WT and all population as described in Section 6.2.2 in main SAP v1.0 dated 17 MAY 2019.
2) The analyses of ORR for confirmed FGFR positive subjects in WT and all analysis population at the sponsor's discretion, as described in Section 6.2.2 in main SAP v1.0 dated 17 MAY 2019.
3) A sensitivity analysis will be performed by only including FGFR DNA-positive patients.

Mutant subgroup:

Two sensitivity analyses will be performed for ORR of mutant subgroup as described in Section 6.2.2 in main SAP v1.0 dated 17 MAY 2019.

- First analysis will include all subjects with both known or unknown PIK3CA and RAS mutation status, per IxRS randomization scheme.

- Second analysis will only include subjects with known PIK3CA and RAS mutant, per IxRS randomization scheme.

### 6.2.2 Secondary Efficacy Analysis

#### 6.2.2.1 Changes in Secondary efficacy endpoints

ORR is considered as primary efficacy endpoint for Phase 2 part. The remaining secondary efficacy endpoints include PFS, DCR, and DOR and will be analyzed as described in main SAP v1.0 dated 17 MAY 2019.

#### 6.2.2.2 Analysis of secondary efficacy endpoints

The secondary efficacy variables based on final database will be tested in the WT subgroup and all study population. No step-wise hypothesis test will be performed.

**Analysis of PFS**

Refer to main SAP v1.0 dated 17 MAY 2019

**Analysis of DCR**

Refer to main SAP v1.0 dated 17 MAY 2019

**Analysis of DOR**

Refer to main SAP v1.0 dated 17 MAY 2019

#### 6.2.2.3 Sensitivity Analysis of Secondary Efficacy Endpoints

Refer to main SAP v1.0 dated 17 MAY 2019 for analyses of PFS, DCR and DOR.

Changes in sensitivity analyses for mutant patients:

Two supportive analyses were originally proposed for PIK3CA and RAS mutant patients: First analysis will be performed by including mutant patients and those with unknown mutation status, per IxRS randomization scheme; second analysis will only include known PIK3CA and RAS mutant patients.

However, due to very small number of enrolled subjects with known mutant status at randomization, for time to event endpoints (PFS and DOR), the supportive analyses will only be performed by combing mutant patients and those with unknown mutation status, per IxRS randomization scheme.

### 6.2.3 Changes in Statistical test strategy for secondary endpoints

The proposed sequential test strategy in main SAP will not be applied.

### 6.2.4 Exploratory Efficacy Analysis

**Changes in analysis of OS**

The Step-wise hypothesis test and analysis of overall survival proposed in the main SAP will not be performed. OS will be analyzed in a similar way to PFS for WT and all population.

### 6.2.5 Subgroup Analyses

Subgroup analyses will be performed in a descriptive fashion for ORR, DCR, PFS and OS. For ORR and DCR, subgroup analysis will only be performed within the subgroup which has at least 10 subjects in total across two arms. For PFS and OS, subgroup analysis will only be provided if there is at least 10 events in total within the subgroups across two arms. Both central and investigator assessment will be used for ORR, DCR and PFS.

Subgroup levels to be analyzed are:

- Stratification factors at randomization (according to IxRS); However, in case of significant discrepancies between stratification factors entered in the interactive voice/web response system (IxRS) and information entered in CRF, the information from IxRS will be used for primary analysis but information from CRF will also be summarized
- Region (Asia vs. North America [NA]/Western Europe [W EU]/Israel/Australia vs. Rest of World [RoW])
- Age (<65 vs. ≥ 65, calculated at the date of randomization using date of birth)
- ECOG performance status (0 vs. 1)
- Gender (male vs. female)
- Race
- Renal function (GFR <45 vs. 45 - <60 vs. ≥60 mL/min/1.73m²)
- hepatic function (AST and ALT each categorized by ≤ULN vs. >ULN ULN)

## 6.3 Safety

### 6.3.1 Adverse events (AEs)

Refer to main SAP v1.0 dated 17 MAY 2019

New analysis added:

The summary of TEAEs will be provided for following subjects:

- For subjects with FGFR positive reading at enrollment but considered negative by confirmation reading.

- For subjects with FGFR score that went from 3+/4+ by original reading at enrollment to 0 by confirmation reading
- For subjects with FGFR score that went from 3+/4+ by original reading at enrollment to 0/1/2 by confirmation reading

**Adverse Events of Special Interest**

Refer to main SAP v1.0 dated 17 MAY 2019.

### 6.3.2 Clinical laboratory evaluations

Refer to main SAP v1.0 dated 17 MAY 2019

New analysis added:

Summary of subjects with >ULN phosphate will be provided by treatment arm.

### 6.3.3 Other safety measures

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.3.4 Pharmacokinetics/pharmacodynamics

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.3.5 Biomarker evaluation

Refer to main SAP v1.0 dated 17 MAY 2019.

### 6.3.6 Changes in Patient-reported outcomes (PRO)

An analysis of covariance (ANCOVA) as well as other sensitivity analysis proposed in main SAP will not be performed.

## 7. Document history and changes in the planned statistical analysis

1. Main SAP v1.0 dated 17 MAY 2019

## 8. References

Not applicable.

# Title page

## A randomized, open label, multicenter Phase 2/3 study to evaluate the efficacy and safety of rogaratinib (BAY 1163877) compared to chemotherapy in patients with FGFR-positive locally advanced or metastatic urothelial carcinoma who have received prior platinum-containing chemotherapy

**Phase 2/3 study of rogaratinib (BAY 1163877) vs chemotherapy in patients with FGFR-positive locally advanced or metastatic urothelial carcinoma**

| | | | |
|---|---|---|---|
| Bayer study drug | BAY 1163877 / rogaratinib | | |
| Study purpose: | Assess the efficacy and safety of rogaratinib compared to chemotherapy | | |
| Clinical study phase: | Phase 2/3 | Date: | 19 AUG 2020 |
| Study No.: | 17403 | Version: | 2.0 |
| Author: | PPD | | |



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

## Abbreviations

| | |
|---|---|
| AE | Adverse event |
| CSR | Clinical Study Report |
| PRO | Patient-Reported outcome |
| SAP | Statistical Analysis Plan |

# 1. Introduction

This Supplemental SAP describes analyses that were not included in the main SAP version 1.0 (dated 12 NOV 2019) and SAP supplement version 1.0 (dated 24 APR 2020) but may be used for CSR.

# 2. Study Objectives

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 (dated 24 APR 2020).

# 3. Study Design

Refer to the Integrated Protocol v3.0 (dated 12 NOV 2019) Section 5 for study design.

## 3.1 Study periods and duration

Refer to main SAP v1.0 dated 17 MAY 2019.

## 3.2 Schedule of procedures

Refer to the Integrated Protocol v3.0 (dated 12 NOV 2019) Table 9-1 for the detailed schedule for the assessments.

## 3.3 Determination of sample size

Refer to the Integrated Protocol v3.0 Section 10.4 (dated 12 NOV 2019) for the details about sample size calculation.

# 4. General Statistical Considerations

## 4.1 General Principles

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.1.1 Population characteristics and safety

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 (dated 24 APR 2020).

## 4.2 Handling of Dropouts

Refer to main SAP v1.0 dated 17 MAY 2019

## 4.3 Handling of Missing Data

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.3.1 Time-to-event variables

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.3.2 Response rates

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.3.3 Adverse Event Start Date

Refer to main SAP v1.0 dated 17 MAY 2019

## 4.4 Interim Analyses and Data Monitoring

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 (dated 24 APR 2020).

## 4.5 Data Rules

### 4.5.1 Time intervals

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.5.2 Baseline

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.5.3 Repeated measures

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.5.4 Stratification

Refer to main SAP v1.0 dated 17 MAY 2019

### 4.5.5 Region

Refer to main SAP v1.0 dated 17 MAY 2019

## 4.6 Validity Review

Refer to main SAP v1.0 dated 17 MAY 2019

# 5. Analysis Sets

Refer to main SAP v1.0 dated 17 MAY 2019

# 6. Statistical Methodology

## 6.1 Population characteristics

### 6.1.1 Disposition

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.1.2 Protocol deviation

Refer to main SAP v1.0 dated 17 MAY 2019

New analysis added:

Only three subjects were impacted by Covid-19 for this study. The protocol deviations associated with Covid-19 will be listed.

### 6.1.3 Demographics and other baseline characteristics

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.1.4 Medical history

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.1.5 Prior and concomitant medication, Prior local Anti-cancer therapy, Systemic Anti-Cancer Therapy during follow-up, Diagnostic and Therapeutic procedures, and Radiotherapy

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.1.6 Extent of Exposure

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 dated 24 APR 2020.

## 6.2 Efficacy

### 6.2.1 Primary Efficacy Variable

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 dated 24 APR 2020.

### 6.2.2 Secondary Efficacy Analysis

#### 6.2.2.1 Secondary efficacy endpoints

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 dated 24 APR 2020.

#### 6.2.2.2 Sensitivity Analysis of Secondary Efficacy Endpoints

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 dated 24 APR 2020.

### 6.2.3 Statistical test strategy for secondary endpoints

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 dated 24 APR 2020.

### 6.2.4 Exploratory Efficacy Analysis

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 dated 24 APR 2020.

### 6.2.5 Subgroup Analyses

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 dated 24 APR 2020.

## 6.3 Safety

### 6.3.1 Adverse events (AEs)

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 dated 24 APR 2020.

**Adverse Events of Special Interest**

Refer to main SAP v1.0 dated 17 MAY 2019.

### 6.3.2 Clinical laboratory evaluations

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 dated 24 APR 2020.

### 6.3.3 Other safety measures

Refer to main SAP v1.0 dated 17 MAY 2019

### 6.3.4 Changes in Pharmacokinetics/pharmacodynamics

Old text in main SAP dated 17 MAY 2019:

The details of pharmacokinetic analysis and results will be described in the separate analysis plan and report.

New analysis added:

Plasma samples for measurement of rogaratinib (BAY 1163877) concentrations will be collected from all rogaratinib-treated patients as follows:

- Day 1 of Cycle 1, 2, 3, 4 and 5:
    - at pre-dose (before supervised dose administration)
    - 1 (±0.5h) hour post-dose (between 0.5 and 1.5 after dose administration)

In the clinical study report, only plasma concentration data of individual subject for all analytes will be listed.

### 6.3.5 Biomarker evaluation

Refer to main SAP v1.0 dated 17 MAY 2019.

### 6.3.6 Patient-reported outcomes (PRO)

Refer to main SAP v1.0 dated 17 MAY 2019 and SAP supplement version 1.0 dated 24 APR 2020.

## 7. Document history and changes in the planned statistical analysis

1. Main SAP v1.0 dated 17 MAY 2019
2. SAP supplement v1.0 dated 24 APR 2020.

## 8. References

Not applicable.