STATISTICAL ANALYSIS PLAN

Adjuvant Axitinib Treatment of Renal Cancer: A Randomized Double-blind Phase 3 Study of Adjuvant Axitinib vs. Placebo in Subjects at High Risk of Recurrent RCC

Study No. AP311736

Authorization Signature Page

Adjuvant Axitinib Treatment of Renal Cell Cancer: A Randomized Double-blind Phase 3 Study of Adjuvant Axitinib vs. Placebo in Subjects at High Risk of Recurrent RCC

Author:

Lynya Engel

06MAR18

Lynya Engel PhD
Statistical Scientist, Biostatistics
IQVIA
Date

Approved by:

Rolland Gaudet

06MAR2018

Rolland Gaudet
Director, Biostatistics
IQVIA
Date

Sponsor Representative:

06. March 2018

Rolf Linke, MD
CMO
SFJ Pharma II SPC, Ltd.
Date

Table of Contents

Modification History

Unique Identifier for this Version	Date of the Document Version	Author	Significant Changes from Previous Authorized Version
Draft 1.0	05DEC2011	Mark Shaw	N/A – First Version
Draft 1.1	10JAN2012	Mark Shaw	Updates based on customer comments and updated version of the protocol (Version 3)
Draft 2.0	11JAN2012	Mark Shaw	Updates following Senior Biostatistical Review of Version 1.1
Final 1.0	23JAN2012	Mark Shaw	Updates based on customer comments on Draft V2.0
Final 2.0	27JAN2012	Mark Shaw	Update to definition of Signs and Symptoms in Section 8.4
Final 3.0	01MAR2012	Mark Shaw	Updates following customer proposals
Final 4.0	16JUL2012	Mark Shaw	Updates based on Version 4 of the protocol
Final 5.0	31JAN2013	Mark Shaw	Addition of sensitivity analyses based on SFDA comments. Updates based on Version 6 of the protocol. Update to stratification defined in Section 7.2.
Final 6.0	02SEP2013	Caroline O'Brien	Updated based on Version 7 of the protocol.
Draft 7.0	23APR2014	Denise Williamson	Updated based on Version 8 of the protocol.
Draft 8.0	05MAR2015	Troy Johnson	Updated based on Version 9 of the protocol.

Draft 9.0	04JUN2015	Troy Johnson	Feedback from reviewers incorporated
Draft 10.0	22NOV2016	Lynya Engel	Addition of sensitivity analysis including histologically confirmed subjects. Addition of sensitivity analysis of overall survival censoring unblinded subjects at the date of unblinding. Addition of risk group analyses. Renamed subset analyses of all efficacy and safety by country to "Post hoc analyses."
Draft 11.0	JAN2017	Lynya Engel	Changed SAP template. Added Treatment Misallocations section. Added section for testing model assumptions. Added possible exploratory endpoints including RMST. Added censoring rule tables as Appendix A.
Draft 12.0	FEB2017	Lynya Engel	Added futility analysis Added removing a stratification factor in the analysis of the primary endpoint if the model does not

			converge Clarified units of measurement for DFS and OS analyses Added new censoring rule to include censoring subjects who had two or more missed visits Added specification to DFS sensitivity analyses regarding if analyses would be based on IRC Review or Investigator Assessments Added more details to DFS sensitivity analysis using scheduled visits Added more details regarding interval censoring used in parametric model Removed DFS analyses stratified by risk group only Removed DFS analyses stratified by country only Removed DFS analysis by years of treatment Changed model for subgroup analyses to not include stratification factors

			Changed what had previously been defined as "subgroup" to "risk group" Removed Overall Survival sensitivity analyses Modified summaries of dose interruptions, dose reductions and dose increases Added guidelines for imputing dates for adverse events Added section on missing data in safety endpoints Defined Day 1 for laboratory analyses and expanded details for worst category summarizations
Draft 13.0	FEB2017	Lynya Engel	Added details for how to calculate discordance rates Methods for evaluating the validity of model assumptions for overall survival were removed Added clarity for how baseline will be defined for laboratory values Removed shift tables demonstrating frequencies of ECOG increases of ≥ +1 and +2 change from baseline to worst value after first dose Removed summaries of subjects

			with a Worst-Grade of >=Grade 4 TEAE and summaries of subjects with a Worst-Grade of >=Grade 4 related TEAE Removed summaries of subjects with a Grade 5 TEAE through 28 days of last dose of the study treatment judged not to be causally related to PD Removed summaries of subjects with a Grade 5 TEAE > 28 days after last dose of the study treatment not judged to be causally related to PD Added details for start date of anti-tumor treatments during follow-up
Draft 13.1	MAR2017	Lynya Engel	Updated name of Sponsor Representative
Draft 14.0	APR2017	Lynya Engel	Updated based on Version 10 of the protocol Updated list of abbreviations Added analysis of relative dose Deleted treatment compliance analyses Added exploratory landmark analysis Revised Appendix A

			Added Appendix B
Draft 15.0	APR2017	Lynya Engel	Updates based on customer comments
Final 7.0	APR2017	Lynya Engel	Updates based on customer comments
Final 8.0 (draft v1.0)	DEC2017	Lynya Engel	Updates based on customer comments
Final 8.0 (draft v2.0)	FEB2018	Lynya Engel	Added clarification to censoring rules and stopping criteria.

List of Abbreviations and Definition of Terms

AE	Adverse Event
AJCC	American Joint Committee on Cancer
ALT	Alanine Aminotransferase
ANC	Absolute Neutrophil Count
AST	Aspartate Aminotransferase
AT	As-Treated
ATC	Anatomical Therapeutic Chemical
BID	Twice a day
BMI	Body Mass Index
BP	Blood Pressure
BUN	Blood Urea Nitrogen
CI	Confidence Interval
CRF	Case Report Form
CT	Computed Tomography
CTCAE	Common Terminology Criteria for Adverse Event
DBP	Diastolic Blood Pressure
DFS	Disease-Free Survival
DMC	Data Monitoring Committee
EDR	Early Discrepancy Rate
ECG	Electrocardiogram
ECOG PS	Eastern Collaborative Oncology Group Performance Status
HLGT	High Level Group Term
HLT	High Level Term
HR	Hazard Ratio
HRAs	Health Regulatory Authorities
IA	Interim Analysis
ICH	International Conference on Harmonisation
INR	International Normalizing Ratio
INV	Investigator Assessment
IRC	Independent Review Committee
IVRS	Interactive Voice Response System
IWRS	Interactive Web Response System
LDH	Lactate Dehydrogenase
LDR	Late Discrepancy Rate
MedDRA	Medical Dictionary for Regulatory Authority
MG	Milligrams
MRI	Magnetic Resonance Imaging

MS	Millimeters/Second
NCI	National Cancer Institute
NLR	Neutrophil to Lymphocyte Ratio
OS	Overall Survival
PH	Proportional Hazards
PI	Principal Investigator
PS	Performance Status
PT	Preferred Term
QTcB	QT interval corrected for heart rate using Bazett's method
QTcF	QT interval corrected for heart rate using Fridericia's method
RCC	Renal Cell Carcinoma
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SBP	Systolic Blood Pressure
SMQ	Standardised MedDRA Query
SOC	System Organ Class
StD	Standard Deviation
TEAE	Treatment Emergent Adverse Event
TSH	Thyroid Stimulating Hormone
ULN	Upper Limit of Normal
UPC	Urine Protein to Creatinine
WBC	White Blood Cell
WHO	World Health Organization
YR	Year

1. Introduction

This document describes the rules and conventions to be used in the presentation and analysis of efficacy and safety data for protocol AP311736. It describes, in detail, the data to be summarized and analyzed, including specifics of the statistical analyses to be performed.

This statistical analysis plan is based on the final Protocol version 10[1], dated 21March 2017. This document may modify the plans outlined in the protocol and supersedes the protocol; however, any major modifications of the primary endpoint or its analyses will be reflected in a protocol amendment.

Any deviations from this analysis plan will be described in the Clinical Study Report.

2. Study Objectives

2.1. Primary Objective

To demonstrate an improvement in disease free survival (DFS) in subjects at high risk of recurrent renal cell carcinoma (RCC) randomly assigned to adjuvant axitinib vs. placebo after nephrectomy.

2.2. Secondary Objectives

- Compare overall survival (OS) associated with axitinib to that associated with placebo;
- Assess safety/toxicity profile of administration of axitinib

3. Study Design

3.1. General Description

This is a prospective, randomized, double-blind placebo controlled Phase 3 trial of oral axitinib starting at 5 mg twice daily vs. placebo given for up to 3 years, with a minimum of 1 year, in subjects at high risk of recurrent RCC aged 18 or over (20 or over in Japan, Korea and Taiwan). The dose may be increased or decreased depending on individual subject tolerance of axitinib.

Approximately 700 subjects will be randomized in a 1:1 ratio between axitinib vs placebo. Subjects should begin study treatment within 7 days after randomization. Randomization will occur no sooner than 4 weeks post nephrectomy and no later than 12 weeks post nephrectomy.

Recurrence or occurrence of a secondary malignancy will be followed up at clinic visits until the time of the final analysis (regardless of the duration of treatment) every 16 weeks for the first 3 years from initiation of study treatment (also, at the end of treatment) and every 6 months

thereafter for the remainder of the follow-up period. The total duration of the study from first subject in to final analysis is estimated to be 5 years.

All subjects will be followed for OS status (regardless of the duration of treatment) every 16 weeks until the time of final DFS analysis. Survival status can be ascertained by telephone contact. At the time of final DFS analysis, OS results will be analyzed and summarized.

A single interim analysis (IA) of efficacy and safety will be performed after approximately 184 DFS events (approximately 75% of the total number of required events as assessed by the independent review committee [IRC]) have occurred. If the event occurrence pace is much slower than anticipated, the IA could be performed when less than 184 events are observed (<75% of the total required) at the sponsor's discretion with the Data Monitoring Committee's (DMC) consensus. To protect the integrity of the study and to preserve the Type 1 error, a fraction of alpha will be spent at the interim analysis of DFS based on an O'Brien-Fleming spending function. The stopping boundaries will be calculated based on the actual number of events at the time of the interim analysis.

The objectives of the interim analysis will be:

- To assess the safety, including any unexpected toxicity.
- To allow for early stopping of the trial due to futility.
- To assess the efficacy of the study drug to allow stopping of trial for success of efficacy.

3.2. Treatments

3.2.1. Treatments Administered

Subjects will receive axitinib (at a starting dose of 5 mg twice a day [BID]) or placebo, identical in appearance to the active study drug, BID. Both should be taken orally with or without food.

Dose adjustments, including dose increase or dose reduction will be based on adverse events (AEs) experienced by the individual subject. Study drug will be taken beginning on Day 1 of the study. Doses should be taken approximately 12 hours apart for continuous dosing. Subjects should be instructed to take their doses at approximately the same times each day. If a subject vomits or misses a dose, an additional dose should not be taken. The next prescribed dose should be taken at the usual time. Missed or vomited doses, must be indicated in the source documents and case report forms (CRFs). Study treatment will be administered in cycles of 4 weeks in duration.

Subjects who tolerate axitinib or placebo with no AEs related to study drug above Common Terminology for Adverse Events (CTCAE) Grade 2 for a consecutive 2-week period may have their dose increased by one dose level to maximum of 10 mg BID as per Table 1 (unless the subject's blood pressure [BP] is >150/90 mm Hg or the subject is receiving antihypertensive medication). The clinical judgment of the treating physician should be exercised in titrating the

axitinib/placebo dose.

Subjects experiencing adverse drug reaction should undergo dose modification as recommended in Section 5.3.4 of the protocol.

Once the dosage is reduced, it can be uptitrated again.

Concomitant medications that are known to substantially inhibit the enzyme, CYP3A4/5, should be avoided as much as medically possible on this study as in Section 5.5.1 of the protocol. If a strong CYP33A4/5 inhibitor must be co-administered, based on investigator judgment, the dose of study drug (axitinib/placebo) should be decreased by one or more dose levels.

Table 1 Available Axitinib/Placebo Dose Levels

Dose Level	Dose	Dispensed As
+2	10 mg BID	2 X 5 mg Tablets BID
+1	7 mg BID	1 X 5 mg Tablet BID + 2 X 1mg Tablets BID
0 (Starting Dose)	**5 mg BID**	**1 X 5 mg Tablet BID (twice daily)**
-1	3 mg BID	3 X 1 mg Tablets BID
-2	2 mg BID	2 X 1 mg Tablets BID
-3	1 mg BID	1 X 1 mg Tablet BID

3.2.2. Method of Assigning Subjects to Treatment Groups

Subjects must be randomized no earlier than 4 weeks and no later than 12 weeks after nephrectomy and treatment should be started within 7 days after randomization. Subject eligibility must be confirmed by IRC assessment of imaging by a sponsor designated center prior to randomization.

Subjects must be randomized based on the assessment by the IRC when there is a discrepancy between the local and the IRC imaging review.

A centralized system will be used to assign Subject numbers and randomize subjects to blinded study drug: axitinib, or blinded placebo identical in appearance to the active study drug.

After a subject has provided written informed consent and has completed the necessary screening assessments, the clinical site must contact a centralized internet/telephone registration system (IWRS/IVRS), to enroll the subject into study.

At the time of registration, the clinical site staff must provide site and subject identifiers and demographic information. The registration system will assign a unique subject identification number. The system will also be used to assign blinded study medication bottles. Subjects will be randomized to one of two arms:

Axitinib 5 mg BID taken orally or Placebo tablets BID taken orally and will be stratified by;

1. Risk group*
    a. pT2, pN0 or pNx, M0 and ECOG PS 0-1
    b. pT3, pN0 or pNx, M0 and ECOG PS 0-1
    c. pT4, pN0 or pNx, M0 and ECOG PS 0-1
    d. Any pT, pN1, M0 and ECOG PS 0-1

* Risk Groups are based on American Joint Committee on Cancer (AJCC) TNM staging version 2010 and Eastern Collaborative Oncology Group (ECOG) performance status (PS)

2. Country (note only those subjects in risk group a or b will be stratified by country)
    a. Japan
    b. China (Mainland)
    c. Korea
    d. Taiwan
    e. Hong Kong
    f. India
    g. USA
    h. France
    i. Spain

Eligible subjects will be initially randomized in a 1:1 ratio to one of the two treatments using the registration system.

Randomization shall continue until a minimum of 10% of subjects are represented in this study from risk group c and d. Based on amendment 8 of the protocol, risk group a was closed to

further enrollment, to keep the number of subjects in risk group b, c, and d to about 90% or more.

3.2.3. Blinding

The study is double-blind. Axitinib will be supplied as 1 mg and 5 mg film-coated tablets for oral administration in light-protecting bottles. Placebo will match all dose formulations and will be identical in appearance to the active study drug.

At the initiation of the trial, the trial site will be instructed on the method for breaking the blind. The method will be an electronic process. Blinding codes should only be broken in emergency situations for reasons of subject safety, or if the subject has a confirmed recurrence or occurrence of a secondary malignancy with limited alternative treatment option and knowledge of study drug is required to facilitate further treatment decisions. For those with suspected recurrence, the site must await results from the IRC. The decision to break the blind must be approved by the Sponsor prior to doing so. When the blinding code is broken, the reason must be fully documented in the site source document. In these instances, only the principal investigator (PI) is unblinded. The broader study team will remain blinded.

Where unblinded study data is to be produced for review by the DMC for the study, a biostatistics team in a separate location from the study team will be used to produce and distribute the unblinded data. The data will only be distributed to unblinded personnel as agreed in a separate DMC analysis plan and/or DMC charter.

3.3. Determination of Sample Size

The subject population in this study can be classified into 4 risk groups, as defined in section 3.2.2.

Sample size was determined based on the analysis on the primary endpoint, DFS.

The sample size for this study was calculated based on the following assumptions:

- Time to DFS event follows an exponential distribution.
- The percentage of subjects randomized from the 4 risk groups, 2-year DFS rates for placebo arm and axitinib arm are assumed below:

Table 2 DFS Rates Per Risk Group

Categories	Risk groups	Percentage of subjects	2-year DFS rate for placebo arm	2-year DFS rate for axitinib arm
1	a. pT2/pN0 or pNX/M0 b. pT3/pN0 or pNX/M0	90%	70%	79%
2	c. pT4/pN0 or pNX/M0	1%	35%	61%
3	d. Any pT/pN1/M0	9%	33%	51%

The assumptions of 2-year DFS rates for the placebo arm and axitinib arm are equivalent to the assumptions of hazard ratios to be 0.66, 0.47, and 0.61 for the 3 categories 1, 2 and 3 respectively. Given the assumed distribution of subjects randomized (1:1) into each of these categories, the overall hazard ratio is estimated to be 0.654.

Based on the above assumptions, a minimal number of 245 DFS events will be required to provide 90% power to detect a hazard ratio (HR) of 0.654 between the two treatment groups with 2-sided significance level of 0.05. The nominal significance level assuming a single interim (at 75% of the required events) and a final analysis for efficacy will be determined by using the Lan-DeMets procedure with an O'Brien-Fleming stopping rule. At the final analysis, an observed HR $\leq$0.773 would be required for statistical significance with the overall Type I error rate preserved at the nominal 0.05 level.

Applying a 1:1 randomization and a planned accrual period of 24 months, a maximum study period of 60 months (5 years), it was estimated that approximately 700 subjects will be required. This assumes a 23% drop-out rate by 18 months.

The final analysis will take place when approximately 245 DFS events are observed.

3.4. Changes in the Conduct of the Study or Planned Analyses

A summary of the key amendments that occurred between November 2011 and March 2017 with an impact on the conduct of the study and planned analyses are described below.

3.4.1. Changes in the Conduct of the Study

Amendment#6 (Jan 2013): (a) inclusion criteria was modified and sample size determination had been adjusted to include any Furhrman grade and the subject population was classified into 4 risk groups: (1) pT2, pN0 or pNx, M0 and ECOG PS 0-1; (2) pT3, pN0 or pNx, M0 and ECOG PS 0-1(3) pT4, pN0 or pNx, M0 and ECOG PS 0-1 (4) Any pT, pN1, M0 and ECOG PS 0-1.

Amendment#7 (Aug 2013) A change was made that all subjects would be followed up for OS status every 16 weeks until the time for final DFS analysis rather than until the time for analysis of OS. Instructions were added that tumor imaging should be obtained if they were not performed within 8 weeks of the end of study or at withdrawal for subjects who discontinue treatment for reasons other than disease progression. A modification was made that randomization would continue until a minimum of 10% of subjects are represented in the study from risk groups c and d. Other modifications included allowing unscheduled assessments to occur coincident with safety events, early discontinuation of study treatment, early study termination, suspected recurrence or suspected occurrence of a secondary malignancy between protocol specified study visits. Instructions were added that for subjects who present with findings suggestive of a tumor recurrence or secondary malignancy, histopathological confirmation of the diagnosis should be obtained except in subjects whose lesions are deemed by the investigator not to be amenable to

biopsy. For these subjects, recurrent disease would be documented using Magnetic Resonance Imaging (MRI) or Computed Tomography (CT) if MRI is not available or is contraindicated. In addition, a modification was made that all histopathologically confirmed secondary malignancy must be reported as a Serious Adverse Event (SAE).

Amendment#8 (Feb 2014) (a) Inclusion criteria was modified to increase the number of subjects with higher risk of RCC recurrence: subjects with pT2, pN0 or pNx, M0 and ECOG PS 0-1 were excluded from enrollment in the trial. (b) Assuming a 23% drop-out rate, the sample size has been adjusted from 592 subjects to 692 enrolled subjects to account for the drop-out.

3.4.2. Changes from Analyses Planned in the Protocol

Amendment#6 (Jan 2013): - New sensitivity analyses were added to analysis of primary endpoint of DFS as described in Section 7.3.1.

Amendment#9 (Oct 2014)- The interim analysis plan was changed from two initial planned interim analyses (one at 47 events or 20% of the events as assessed by IRC and one at 142 events) to one interim analysis that will take place at 184 events (75% of events as assessed by IRC). To protect the integrity of the study and to preserve the Type 1 error, a fraction of alpha will be spent at the interim analysis based on an O'Brien Fleming spending function. The objectives of the new planned interim analysis will be: (a) to assess the safety, including any unexpected toxicity; (b) to allow for early stopping of the trial due to futility; (c) to assess the efficacy of the study drug to allow stopping of trial for success of efficacy.

Amendment#10 (Mar 2017): The primary endpoint was modified to include evaluations of available local histo-/cytopathology reports by an independent IRC oncologist for cases in which recurrence or occurrence of secondary malignancy had been confirmed by the site in the absence of IRC imaging confirmation. Specifically, the date of recurrence or the occurrence of secondary malignancy will be defined as the date of the tumor scan or the date of collection of the histo-cytopathological specimen (for subjects who have site confirmed recurrence or occurrence of secondary malignancy in the absence of IRC imaging confirmation) that demonstrated unequivocal recurrence or secondary malignancy according to protocol criteria. If both imaging and histo-/cytopathological confirmation of recurrence or secondary malignancy are available, the earlier of the two dates will be considered. In addition, the protocol was amended to include censoring subjects who have two or more consecutively missed or not readable scans immediately prior to an event. The date of censoring in these instances will be on the date of the most recent scan prior to the missing/not readable scans. The protocol was further amended to remove excluding as an event the occurrence of a second primary cancer that was basal cell carcinoma, squamous cell skin cancer or in situ carcinoma of the cervix uteri. Additionally, the protocol was also changed to record anti-tumor therapy even after 28 days after last treatment. More details were provided for analysis of DFS defined by assigning dates for events and censoring at scheduled scan dates.

4. Analysis Populations

Screen failure subjects will not be included in any analyses.

Subjects who start any of the study drugs (i.e., axitinib or placebo) will be considered to have started treatment.

4.1. Intent-To-Treat Population

Intent-to-Treat (ITT) Population: This population will include all randomized subjects regardless of whether or not treatment was administered and will be based on randomized treatment assignment. This population will be the primary population for evaluating baseline characteristics and efficacy.

4.2. As-Treated Population

As-Treated (AT) Population: The As-Treated population consists of all subjects who received at least 1 dose of study medication with treatment assignments designated according to actual study treatment received. This population will be the primary population for evaluating treatment administration/compliance and safety.

The number and percentage of subjects in each population will be summarized by treatment group and listed.

4.3. Treatment Misallocations

If subjects were:

- Randomized but not treated, then they will be reported under their randomized treatment group for efficacy analyses. However, they are by definition excluded from the safety analyses as actual treatment is missing.

- Randomized but took incorrect treatment, then they will be reported under their randomized treatment group for all efficacy analyses, but will be reported under the treatment they actually received for all safety analyses.

A summary table of treatment misallocations will be provided.

4.4. Protocol Deviations

Subjects with major protocol deviations will be summarized by treatment group.

Major protocol deviations include, but are not limited to:

- Subjects who did not meet inclusion/exclusion criteria or eligibility was not adequately verified
- Received prohibited concomitant medications post-baseline
- Developed withdrawal criteria but were not withdrawn
- Has a dosing error likely to impact key outcome measures (for example subject did not receive the medication they were randomized to receive)

5. Efficacy Endpoints and Covariates

5.1. Primary Efficacy Endpoint

The primary endpoint is DFS, defined as the time interval (measured in years) from the date of randomization to the first date of recurrence (distant or local recurrence of RCC) or the occurrence of a secondary malignancy or death due to any cause. The primary DFS analysis will be based on assessment by the IRC. The IRC will review all available scans and in cases for which there is site confirmation of recurrence or occurrence of a secondary malignancy in the absence of IRC imaging confirmation, available local histo-/cytopathology reports will be reviewed by an independent IRC oncologist.

Recurrence refers to relapse of the primary tumor in situ or at metastatic sites. The date of recurrence or the occurrence of a secondary malignancy is defined as the earlier of:

- the date of the tumor scan or the date of collection of the histo-/cytopathological specimen (for subjects who have site confirmed recurrence or occurrence of a secondary malignancy in the absence of IRC imaging confirmation) that demonstrated unequivocal recurrence or a second malignancy according to protocol criteria

Per the IRC charter, radiological findings should be unequivocal. At the time of the final IRC review for a given data cut (interim or final), equivocal disease, identified by radiology requiring confirmation, is noted in the IRC data at the 'Global Review' with an indication of disease present and a date of recurrence/secondary malignancy recorded by the IRC as not applicable (N/A). As these findings require confirmation, they will not be counted as events in the DFS analysis unless, there is a subsequent pathology report which provides confirmation of the disease seen on the scan. In this case the scan date should be used as the date of recurrence or secondary malignancy.

In the absence of pathology to confirm the equivocal findings, these subjects will be censored at the date of last scan before anti-cancer therapy (if applicable). The last scan date includes the scan where the equivocal finding was noted.

For subjects identified as having the presence of disease at baseline by IRC review,

- Recurrence will be considered to have occurred on the date of randomization.

For subjects with no DFS event, DFS time will be censored

- at the date of last IRC reviewed scan prior to the time of the final analysis.
- Subjects alive who do not have any post-baseline disease assessments will have their DFS times censored at randomization.

For any subject who receives anti-tumor therapy,

- DFS will be censored on the date of the last IRC reviewed scan prior to taking the anti-tumor therapy. However, in cases where there is no scan prior to anti-cancer therapy medication that was received after randomization, subjects will be censored at the date of randomization.

For subjects who had two or more consecutively missed or not readable IRC reviewed scans immediately prior to a recurrence or occurrence of a secondary malignancy or death,

- DFS will be censored on the date of the last IRC reviewed scan prior to the consecutively missed or not readable IRC reviewed scans.

The length of DFS will be calculated as follows:

DFS (years) = [date of recurrence, secondary malignancy, or death or censor date – randomization date + 1]/365.25.

5.2. Secondary Efficacy Endpoint

Overall survival (OS) is defined as the time from the date of randomization to the date of death due to any cause. In the absence of confirmation of death, survival time will be censored at the last date the subject is known to be alive. Subjects lacking data beyond randomization will have their survival times censored at randomization.

The length of OS will be calculated as follows:

OS (years) = [death date or last known alive date – randomization date + 1] /365.25.

5.3. Safety Endpoints

Adverse Events

Assessment of adverse events will include: type, incidence, severity (graded by the National Cancer Institute [NCI] CTCAE, Version 4.03), timing, seriousness, and relatedness. Laboratory abnormalities will be recorded as an adverse event in accordance with the criteria in Section 8.4 of the protocol.

Baseline tumor-related signs and symptoms will be recorded as adverse events during the trial if they worsen in severity or increase in frequency.

Laboratory Safety Assessments

Hematology and blood chemistry will be drawn and urinalysis performed at the time points described in the Schedule of Activities and analyzed at local laboratories (Refer to Appendix 1 of the protocol). Investigators may have additional blood tests performed for the purpose of planning treatment administration, dose modification, or following adverse events, and these should be repeated as clinically needed. Prothrombin Time/International Normalizing Ratio (INR) should be performed to monitor subjects receiving concomitant warfarin or other anti-coagulants and when clinically indicated.

Thyroid Function Tests (Thyroid Stimulating Hormone [TSH], free T3 and free T4) should be performed for all subjects at baseline (Cycle 1 Day 1 pre-dose). Subsequently, TSH should be done at Cycle 1 Day 15, Cycle 2 Day 1, Cycle 3 Day 1, Cycle 4 Day 1, Cycle 5 Day 1, then every 8 weeks thereafter starting from Cycle 7 Day 1. TSH, free T3 and free T4 should be done on the next visit for any subject who has an elevated TSH in the previous visit. Hypothyroidism should be treated per standard medical practice to maintain euthyroid state with a normal TSH.

Other Safety Assessments

Other safety assessments to be assessed include: physical examinations, 12-lead electrocardiogram (ECG) (recorded at Screening only), height (recorded at Screening only), weight, vital signs (recorded at clinical study visits) and ECOG PS.

5.4. Covariates

5.4.1. Stratification Factors

As described in Section 3.2.2, randomization in this study is stratified by risk groups based on the AJCC TNM staging version 2010, ECOG PS and by country. Analyses for the primary and secondary endpoints as well as sensitivity analyses will be conducted based on analyses stratified by risk group only. Country will not be used as a stratification factor in analyses due to the limited number of subjects anticipated to be enrolled in some countries. Additional supportive analyses will not be stratified.

In the event a subject was randomized to the incorrect strata, the strata used for analysis will be obtained from the IWRS.

6. Statistical Methods

6.1. General Methodology

SAS version 9.2 or higher will be used in the statistical analysis.

The level of significance for the analysis of the primary variable at the interim and final analysis stages are described in Section 6.3. 95% confidence intervals (CIs) will be employed where appropriate

Data for subjects who failed screening will be listed. This may include demographics, inclusion/exclusion criteria responses and reason for screen failure.

6.1.1. Analyses for Continuous Data

For continuous data, the following summary statistics will be presented: n, mean, standard deviation (StD), median, minimum and maximum.

Mean and median values will be reported to one decimal place greater than the original data they were collected from while the StD will be reported to two decimal places greater than the original data. Minimum and maximum values will be reported with the same precision as they were collected.

6.1.2. Analyses for Categorical Data

For categorical variables, statistical summaries will include counts and absolute or relative percentages. The number of missing values will be presented where necessary.

Percentages will be reported to 1 decimal place. Percentages will be calculated using a denominator of all subjects in a specified population.

For by-visit analyses e.g. shift from baseline in toxicity grade of laboratory parameters only subjects with a measurement available will be included in the denominator for that visit.

6.1.3. Analysis of Time-To-Event Endpoints

Time-to-event endpoints will be summarized using the Kaplan-Meier method and displayed graphically when appropriate. Median event times and 2-sided 95% CI for each median will be provided based on the Brookmeyer-Crowley[2] method.

Difference in time-to-event endpoints will be tested using a 2-sided stratified or unstratified log-rank test, where applicable. The stratified analysis will be performed as stated in Section 5.4.1.

Cox proportional hazards model will be used to estimate the hazard ratio and its 95% CI.

A parametric model will also be implemented as a sensitivity analysis of DFS.

6.2. Handling of Dropouts or Missing Data

Subjects who discontinue study treatment will continue to be followed-up for DFS and OS until the time they have an event, or until the time of final analysis for DFS and OS. Subjects who withdraw from the study and are lost to follow-up will not be replaced. Surviving subjects without a DFS event and/or OS event at the time of the final analysis will have their DFS censored as described in Appendix A. All other missing data will be left as missing with the following exception as described in Section 6.2.1.

6.2.1. Missing Data in Adverse Event Dates

Missing dates in adverse events will be imputed as described below:

- For the start date, if the day of the month is missing, the 1st day of the month will be used to replace the missing date unless it is the month of the first dose. In that case, the start date will be imputed as the first dose date. If both day and month are missing, the 1st of January of the non-missing year will be used to replace the missing date unless the non-missing year is the same year as first dose in which case the start date will be imputed as the first dose date.

- For the stop date, if the day of the month is missing, the last day of the month will be used to replace the missing date. If both day and month are missing, December 31 of the non-missing year will be used to replace the missing date.

If the start date is missing for an adverse event, the adverse event is considered to be treatment-emergent.

6.2.2. Missing Data in Safety Endpoints

The percentage of subjects with an adverse event will be calculated using the number of as-treated subjects as the denominator. Therefore, no subjects in the as-treated population are excluded from adverse event displays. The denominator for summary tables for each laboratory parameter will be all subjects in the as-treated population with at least one evaluable cycle for that parameter. Different laboratory parameters may have different denominators, depending on the number of evaluable subjects for each parameter. An evaluable cycle is any cycle with at least one assessment of that parameter. Therefore, subjects with no assessments of a particular laboratory parameter are not included in the analysis of that parameter.

6.3. Interim Analysis and Data Monitoring Check spending function

A single interim analysis (IA) is planned for this study. O'Brien – Fleming type stopping boundaries based on the Lan-DeMets spending function will be applied to the primary endpoint DFS (as assessed by IRC). Futility criteria are not used to calculate the nominal alphas (non-

binding method) in order to control the overall Type I error.

The IA will take place after the first 184 events (75% of planned DFS events as assessed by the IRC) have occurred. If the event occurrence pace is much slower than anticipated, the IA could be performed with less than 184 events observed (<75% of the total required) at the sponsor's discretion with the Data Monitoring Committee's consensus.

The objectives of the IA are:

- To assess safety, including unexpected toxicity.
- To allow for early stopping of the trial due to futility.
- To allow for early stopping of the trial due to efficacy.

If the results of the IA demonstrate statistically significant differences between the 2 treatment arms for DFS (in favor of the active drug), the sponsor, in consultation with HRAs, will disseminate the results of the trial, and the IA may be considered the final analysis for DFS. The nominal level of significance for the interim analysis of DFS determined using the Lan DeMets procedure with an O'Brien Fleming type stopping rule will be 0.0194 (2-sided; if the interim analysis is performed at 184 events).

The descriptive safety analysis as described in Section 7.5 will be presented to assess any toxicities.

The nominal significance level for the interim futility analysis of DFS will be determined using an O'Brien-Fleming stopping boundary. The futility p-value boundary is specified in the table below. If the calculated futility p-value is greater than the pre-specified boundary shown in the table below, the sponsor may choose to stop the trial for futility. The associated critical hazard ratio is also provided in the table below for reference.

Futility Stopping Boundary (non-binding) for DFS for Rejecting Alternative Hypothesis Expressed as HR and p-value

Analysis	Fraction of DFS Events	Number of DFS Events	HR (Axitinib:Placebo)	2-sided p-value
Interim	75%	184	0.836	0.2255

The overall nominal significance level for the efficacy analysis of DFS will be preserved at 0.05 (2-sided test). To protect the integrity of the study and to preserve the type 1 error, a fraction of alpha will be spent at the interim analysis based on an O'Brien-Fleming stopping boundary. The efficacy p-value boundaries are specified in the following table. The associated critical hazard ratio is also provided in the table below for reference.

Efficacy Stopping Boundary for DFS for Rejecting Null Hypothesis Expressed as HRs and p-values

Analysis	Fraction of DFS Events	Number of DFS Events	HR (Axitinib:Placebo)	2-sided p-value
Interim	75%	184	0.708	0.0194
Final	100%	245	0.773	0.0442

The actual nominal α levels for the interim analysis and for the final analysis will depend on the fraction of total events occurred at the time of the IA.

In addition to DFS based on IRC assessment, Kaplan-Meier summaries for DFS (according to investigator assessment) and OS will be presented for each treatment group.

The primary endpoint of DFS is based on IRC assessment. DFS based on investigator assessment is an important sensitivity analysis. Although no formal hypothesis testing is planned for DFS based on investigator assessment, regulatory feedback has indicated the importance of consistency between investigator and IRC assessments of DFS. Therefore, in order to meet the criteria for stopping for efficacy at the time of the interim analysis the same p-value stopping criteria, determined based on the proportion of IRC assessed events at the time of the interim analysis out of the planned 245 events for the final analysis, will be applied to DFS based on investigator assessment. Specifically, both the primary DFS analysis by IRC assessment and the DFS analysis by investigator assessment must meet the stopping criteria to stop for efficacy at the time of the interim analysis. Assuming 184 IRC assessed events at the time of the interim analysis both the IRC and investigator DFS 2-sided p-values will be required to fall below 0.0194 to stop the study for efficacy. The stopping rule for futility will only be applied to the primary DFS analysis by IRC assessment.

Although there is no intention to perform any hypothesis test for OS at the interim analysis, a nominal α of 0.0001 will be allocated to the analysis of OS at the interim. The overall nominal significance level for the efficacy analysis of OS will be preserved at 0.05 (2-sided test).

Note that for the interim analysis, the data will be initially presented to members of the DMC and designated unblinded personnel only. Decisions and recommendations based on the interim analysis review will be disseminated to the sponsor. The sponsor at that time may review and discuss the DMC recommendations with a second panel of independent experts.

There are also planned regular reviews of the safety at separate DMC meetings as needed. Safety data will be summarized as described in Section 7.5. The timing and details of the DMC review are detailed in the DMC charter.

6.4. Multicenter Studies

No adjustment will be made in any analyses for center.

6.5. Multiple Comparisons/Multiplicity

For the interim analysis performed and the final analysis for the primary endpoint of DFS, the significance level has been set using the O'Brien-Fleming method. Although there is no intention to perform any hypothesis test for OS at the interim analysis, a nominal α of 0.0001 will be allocated to the analysis of OS at the interim. There will be no other adjustments for multiple testing.

7. Statistical Analysis

7.1. Analysis of Primary Efficacy Endpoint

7.1.1. Disease-Free Survival Based on IRC Review

Disease-Free Survival of the ITT Population based on the IRC review of tumor assessments will be summarized using the Kaplan-Meier method. Kaplan-Meier curves for each arm will be generated, and median DFS time will also be presented with corresponding 95% CI. The CI for the median will be calculated according to Brookmeyer and Crowley[2]. To test the equality of DFS time in each arm a stratified log-rank test (using a stratification factor as defined in Section 5.4) will be used. Censoring rules as described in Table 1 of Appendix A will be implemented. The statistic of the test will be presented along with the associated two-sided p-value.

The estimated HR (Axitinib/Placebo) will be obtained using a Cox's Proportional Hazards model with treatment group as well as a randomization stratification factor as a covariate. The stratification factor used in the analyses will be risk group as described in Section 5.4.1.

The significance level for the final DFS analysis will be 0.0442 as calculated using the O'Brien-Fleming method. (Assumes interim analysis at 184 events with a significance level of 0.0194).

Annual DFS rates up to 5 years for each treatment group will be estimated and presented with corresponding 95% CI. The CI for the survival function will be calculated using the complementary log-log transformation method.

The final analysis of the primary endpoint of DFS will be performed when approximately 245 DFS events based on IRC review are observed.

A subject listing for DFS will include, DFS time, censoring status, censoring reason, date of randomization and date of last tumor assessment. In addition, a listing of subjects with baseline disease as determined by the IRC, will be produced.

7.1.2. Methods for Evaluating the Validity of Model Assumptions

Schoenfeld residuals for the stratified Cox proportional regression model will be plotted to investigate graphically violations from the proportional hazards (PH) assumption for the primary DFS endpoint; a non-zero slope is evidence of departure from PH. The PH assumption will be formally tested using Schoenfeld's residual test (Schoenfeld, 1980; Therneau and Grambsch, 2000)[3]. Large departures from PH will be evidenced by a p-value < 0.05.

In addition, the proportional hazards assumption will be checked visually by plotting

-log(log(S(t)) versus log(t),

where S(t) is the estimated survival function at time t.

If these assessments show large departures from proportional hazards, then DFS may also be analyzed based on the test of RMST(τ) differences between treatment arms based on the stratified Cox regression model as described in Zhang (2013)[4].

The RMST up to time t* can then be interpreted as the expected survival time restricted to the common follow-up time t* among all patients. Analyses will be repeated using the follow criteria to define t*:

- t*1 = min of (longest observed survival time for experimental arm, longest observed survival time for control arm) in years
- t*2= min of (longest event time for the experimental arm, longest event time for the control arm) in years
- t*3 = the midpoint between the numbers t*1 and t*2

RMST can be estimated consistently by the area under the Kaplan-Meier curve over [0, t*]. The treatment effect between each of the experimental arms and the control arm will be assessed based on the difference in RMST. The associated 95% CI for the difference in means and two-sided p-value will be generated. RMST as a function of t* and the associated treatment effect between each of the experimental arms and the control arm will be plotted against time t*.

Note RMST was not specified in the protocol. This was added to the SAP as an exploratory analysis.

7.2. Analysis of Secondary Efficacy Endpoint

7.2.1. Overall Survival (OS)

OS of the ITT Population will be analyzed in the same way as the primary endpoint. Censoring rules as described in Table 2 of Appendix A will be implemented. At the time of final analysis of DFS, a hypothesis test on OS will be conducted at two-sided alpha = 0.0499 only if DFS is

declared statistically significant.

7.3. Additional Analyses of Efficacy

All additional analyses of efficacy will be performed on the ITT Population unless otherwise specified.

7.3.1. Sensitivity Analyses

The following sensitivity analyses will be performed:

- Investigator Assessed DFS: DFS defined in a similar manner as defined for the primary analysis, but based on local investigator assessment rather than IRC assessment (Refer to Table 3 of Appendix A for censoring rules.) For this analysis tumor imaging assessments as well as histo-/cytopathology information will be used to assess disease recurrence or occurrence of secondary malignancy. If both imaging and histo-/cytopathological confirmation of recurrence or occurrence of secondary malignancy are available, the earlier of the two dates will be considered. In addition, discordance rates between IRC review and investigator assessments using methods described by Amit et al.[5] will be summarized in a table. Calculations will be as follows:

		IRC Review	
Investigator Assessment (INV)		**Recurrence or secondary malignancy or death**	**No recurrence or secondary malignancy or death**
	Recurrence or secondary malignancy or death	a = a1 + a2 + a3	b
	No recurrence or secondary malignancy or death	c	d

a1: number of agreements on timing and occurrence of event;
a2: number of times agreement on event but INV declares event later than IRC Review;

a3: number of times agreement on event but INV declares event earlier than IRC Review;
N= a+b+c+d.

The following measure of discordance will be calculated per arm:

- Total Event Discrepancy Rate: (b+c) / N
- Early Discrepancy Rate (EDR): (a3+b) / (a+b)
- Late Discrepancy Rate (LDR): (a2+c) / (a2+a3+b+c)
- Overall Discrepancy Rate: (a2+a3+b+c) / N

A window of 28 days will be used in determining agreement on timing of an event. Specifically, if the data or recurrence by the IRC is within +/- 28 days of the data of recurrence by the investigator these cases will be considered as 'agreement on the timing and occurrence or an event'.

- DFS based on IRC Without Censoring for New Cancer Therapy or Missed Assessments: DFS based on IRC review, as defined for the primary analysis, but without censoring subjects who:
  - receive further anti-tumor therapy, or
  - have two or more consecutively missed or not readable scans immediately prior to recurrence, or occurrence of a secondary malignancy, or death.

  In other words, a subject who meets the above criteria will be assigned a DFS event at the first date of recurrence or occurrence of a second primary cancer or death. In the absence of a DFS event, DFS time will be censored at the date of last scan prior to the time of analysis. Subjects alive who do not have post-baseline disease assessment will have their DFS times censored at randomization. (Refer to Table 4 of Appendix A for censoring rules.)

- Start of New Anti-tumor Therapy as DFS Event: DFS based on IRC review, as defined for the primary analysis, but utilizing the start date of new anti-tumor therapy as a DFS event. (Refer to Table 5 of Appendix A for censoring rules.)

- DFS at Scheduled Assessments: DFS based on IRC review, as defined for the primary analysis, but instead assigning the dates for events and censoring at the scheduled scan dates instead of the actual scan dates. Specifically, events that occur within +/- 4 weeks of a scheduled scan will be considered an event at the scheduled scan. Events outside the 4-week window will be counted as events at the next scheduled scan time and censoring outside the 4-week window would be censored at the previous scheduled scan. If, however, the event is death or disease at baseline, the date of death or date or randomization respectively will be used as the event date unless anti-tumor therapy was

received prior to the date of death in which case the subject will be censored at the scheduled scan prior to the date of anti-cancer therapy or for subjects who have two or more missed or not readable consecutive tumor scans immediately followed by an event will be censored on the date of their most recent scheduled scan prior to the missing/not readable scans. (Refer to Table 6 of Appendix A for censoring rules.)

- DFS with Interval-censoring: DFS based on IRC review as defined for the primary analysis will be performed with a parametric model utilizing techniques for interval-censored data. Graphical diagnostics will be used to select the distribution of the parametric model that best fits the data. Intervals of 16 weeks will be implemented for censoring purposes (Refer to Table 7 of Appendix A for censoring rules.)

7.3.2. Subgroup Analyses

In addition, DFS by IRC will be analyzed in the same way as the primary analysis for the following subgroups with the exception that there will be no stratification variable in the model due to small numbers of subjects:

- Risk group
  i. a (pT2/pN0 or pNX/M0)
  ii. b (pT3/pN0 or pNX/M0)
  iii. c (pT4/pN0 or pNX/M0)
  iv. d (Any pT/pN1/M0)
  v. c+d combined
  vi. High Risk ([b with Fuhrman grade 3 or 4]+c+d)
  vii. Low Risk (a+[b with Fuhrman grade 1 or 2])
- Age (<65, >=65 years)
- Race
- Gender
- Baseline ECOG PS (0 vs >=1)
- Baseline Weight
  viii. Normal (18.5<=body mass index (BMI)<25)
  ix. Overweight + Obese (BMI>=25)
  x. Overweight (25<=BMI<30)
  xi. Obese (BMI>=30)
- Baseline neutrophil to lymphocyte ratio (NLR) (NLR>3 vs. NLR<=3)
- Fuhrman Grade (1&2 vs. 3 &4)

7.3.3. Post hoc Analyses

In addition to the overall analysis, all efficacy and safety analyses may later be presented by country in order to produce results to be included in Japan, China, Korea and Taiwan sub-population reports. The decision to pursue these analyses will be determined by SFJ depending upon availability of subjects within country strata. There are no other planned subset analyses in this study.

7.3.4. Exploratory Analyses

Exploratory analyses may be performed as deemed appropriate.

RMST for DFS may be explored if non-proportional hazards are observed for the primary endpoint.

DFS and Treatment Duration: A one-year landmark analysis will be performed to explore the association between extended treatment with Axitinib and DFS. Patients randomized to Axitinib will be divided into 2 groups based on treatment status at 1 year. Namely patients whose duration of treatment in less than or equal to one year will be compared to those whose duration of treatment is greater than one year. Only patients randomized to treatment with Axitinib who are alive, disease-free, and still in disease-free survival follow-up at one year will be included in the analysis. Specifically, patients who had disease recurrence, secondary malignancies, died, or were otherwise censored for DFS due to start of new anticancer therapy, missed assessments, or lost to follow-up prior to one year will be excluded from the analysis. Median event times and 2-sided 95% CIs for each median will be provided based on the Brookmeyer-Crowley method. Cox proportional hazards model will be used to estimate the unstratified hazard ratio and its 95% CI. A p-value will not be reported as this analysis is considered exploratory.

Treatment Discontinuations due to AEs of Interest: an evaluation of cumulative incidence of treatment discontinuations for defined AEs of interest will be conducted. A summary table will be provided. A figure displaying the proportion of subjects off treatment over time will be provided for subjects randomized to axitinib or placebo who experienced various defined adverse events of interest.

7.4. Standard Analyses

7.4.1. Disposition of Subjects

Subject disposition, including the number of subjects enrolled, number of subjects treated, number of subjects evaluated for safety and number of subjects who completed the study or withdrew early (including reasons for withdrawal) will be summarized by treatment group for the ITT population.

Disposition information will be listed.

7.4.2. Demographic and Other Baseline Characteristics

The following demographic and baseline characteristics will be summarized for the ITT population.

Demography: Gender, age (calculated from date of birth and date of screening visit), height, weight (at screening), BMI (at screening), ECOG PS (at screening) and race.

Primary Diagnosis: Body Site, time since date of histopathological diagnosis (calculated as number of years from date of histopathological diagnosis to date of screening visit), histological classification, risk group.

Medical History: Medical history will be coded using the latest version of Medical Dictionary for Regulatory Activities (MedDRA) and summarized by system organ class (SOC) and preferred term (PT).

Prior Medication/Therapy: Prior medications, as defined in Section 7.5.5, will be summarized by the Anatomical Therapeutic Chemical (ATC) Classification System and WHO Drug name. Prior non-drug treatments, as defined in Section 7.5.5, will be provided in a listing.

Signs and Symptoms: Signs and symptoms are defined as any events recorded in the medical history page ongoing at screening. They will be summarized by SOC and PT.

7.4.3. Extent of Exposure

Study treatment exposure will be summarized for all subjects in the as-treated population. The number of subjects who completed at least 3 years of study medication, at least 2 years of study medication and at least 1 year of study medication will be provided. Duration of exposure to study medication will be calculated as the number of months from date of first dose to date of last dose + 1 day. For the purposes of categorization of exposure duration, the following conversions (assuming a -14 day window) will be used:

- 1-351 days = "0 - <12 months"
- 352-716 days = "12 - <24 months"
- 717-1081 days = "24 - <36 months"
- >1081 days= "≥36 months".

Reasons for treatment withdrawal will be provided.

Overall duration of exposure to study medication will be summarized by treatment group using summary statistics.

Also, the total dose prescribed in milligrams (mg) and the actual dose administered will be summarized.

Number and percentage of subjects with dose increases, dose reductions and dose interruptions

will be summarized.

Other summaries will include:

- Descriptive statistics for time to the first actual dose interruption, dose reduction and dose increase
- Descriptive statistics for time to the second actual dose interruption, dose reduction and dose increase
- Descriptive statistics for time to the third actual dose interruption, dose reduction and dose increase
- Descriptive statistics for time to first dose interruption, time to first dose interruption that was ≥7 days, ≥ 14 days, and ≥21 days after first dose date. The time to dose interruption is calculated as (start date of the interruption – first dose date + 1).
- Descriptive statistics for time to second dose interruption, time to second dose interruption that was ≥7 days, ≥14 days, and ≥21 days after first dose date.
- Relative dose: percent of actual total dose received relative to intended total dose initially planned per protocol (5mg BID), where actual total dose= total dose received as recorded on CRF, and intended total dose = (prescribed dose at beginning of the study)×(actual dose duration).

7.5. Analysis of Safety

All safety analyses will be performed using the As-Treated population. No formal comparisons between the two treatment arms are planned. Safety and tolerability will be assessed by the incidence of treatment emergent-adverse events (TEAEs), changes in laboratory parameters and vital signs from baseline, and ECG.

An independent DMC will monitor safety during the study on a regular basis. The committee will operate independently from the Sponsor and the clinical investigators.

The primary responsibility of the DMC is to review the accumulating safety data on a regular and an ad hoc basis and make recommendations to the Sponsor regarding the continued conduct of the study. Safety data will be provided at regular intervals to the DMC in the form of summary reports or data listings from the Sponsor or its designated representative.

Details regarding DMC membership, schedule and format of meetings, format for presentation of data, access to interim data, method and timing of providing interim reports to the DMC, and

other issues relevant to committee operations are described in the DMC charter.

The DMC members will use their expertise, experience, and judgment to evaluate the safety data from the trial and to recommend to the Sponsor whether the trial should continue or be stopped early for safety.

7.5.1. Adverse Events

Adverse event (AE) terms recorded on the CRF will be mapped to PT and SOC using the Medical Dictionary for Regulatory Activities (MedDRA). The severity of AEs will be evaluated by the investigator by employing the CTCAE v4.03 (Cancer Terminology Criteria for Adverse Events). The investigator will also judge each event to be "not related" or "related" to study treatment.

A TEAE is defined as any AE with an onset date on or after the date of the first dose of study treatment or any ongoing event on the date of the first dose of study treatment that worsens in severity after the date of the first dose of study treatment.

The AE observation period is defined as the time from date of the first dose of study treatment until 28 days after the last dose of study drug is administered. TEAEs with an onset date through the end of the study observation period will be summarized by treatment arms. The Grade 5 TEAEs with onset date outside of the observation period may be tabulated separately.

For summaries of subject incidence of TEAE, at each level of summarization, a subject will be counted only once for each AE preferred term experienced by the subject within that level (i.e., multiple episodes of events with the same preferred terms will be counted only once). Selected summaries of subject-incidence of AEs (as defined in Section 7.5.1.2) will include the risk difference (defined as the proportion with the event in the axitinib arm minus the proportion with the event in the placebo treatment arm) with 95% CIs and the risk ratio (or relative risk; defined as the proportion with the event in the axitinib group divided by the proportion with the event in the placebo group) with 95% CIs. An overall summary of treatment emergent adverse events (including those that stopped more than 28 days after the decision to discontinue study treatment) will be provided with the number and percent of subjects who experienced the following in each treatment group:

- Subjects with a TEAE
- Subjects with a Related TEAE
- Subjects with a Serious TEAE at any time (including active treatment period plus follow-up period)
- Subjects with a Serious Related TEAE
- Subjects with a Worst-Grade of >=3 TEAE
- Subjects with a Worst-Grade of >=3 Related TEAE
- Subjects with a Grade 5 TEAE at any time

- Subjects with a Grade 5 TEAE through 28 days of last dose of the study treatment
- Subjects with a Related Grade 5 TEAE at any time
- Subjects with a TEAE leading to dose modification (dose reduction or temporary interruption/discontinuation)
- Subjects with a TEAE leading to dose reduction
- Subjects with a TEAE leading to temporary dose interruption/discontinuation
- Subjects with TEAE leading to permanent discontinuation

7.5.1.1 Deaths

All reported subject deaths and whether death was causally associated with the disease under study will be summarized by treatment arm in the as-treated population.

Deaths will be summarized in 2 main categories as follows:

- Deaths within 28 days after the date of receipt of the last dose of study treatment
- Deaths greater than 28 days after the date of receipt of last dose of study treatment

Summary of primary cause of death will be tabulated causality to study disease and relationship to study drug.

7.5.1.2 Treatment Emergent Adverse Events of Special Interest

The following TEAEs of special interests will be summarized for the as-treated population and by SOC and PT for each categorized treatment group. The list of MedDRA defined SOCs and PTs for TEAEs of special interest are provided below and may be updated at the time of database lock based on the MedDRA version in use at the time of reporting and/or any additional safety information available from the Axitinib program that may determine a need to consider additional events of interest. Additionally, for the following special adverse events, Kaplan-Meier methods will be used to display the time to the first occurrence of each of these 13 event categories, if the number of events is >=10 events per treatment arm:

Adverse Events of Special Interest	**Search Terms (MedDRA version 20.1)**
Cardiac disorders	Cardiac disorders – SOC Cardiac and vascular investigations (excluding enzyme tests)- HLGT
Nervous system disorders	Central nervous system vascular disorders- HLGT
Vascular disorders	Embolism and thrombosis- HLGT Pulmonary Embolism PT
Renal vascular disorders	Renal vascular and ischaemic conditions (HLT) Renal failure and impairment (HLT) Renal hypertension and related conditions (HLT)

Haemorrhage	Haemorrhage (excluding lab terms) SMQ (broad and narrow) and Haemorrhage laboratory terms SMQ (broad and narrow)
Hypertension	Hypertension SMQ narrow
Proteinuria	MedDRA PTs: Albumin urine present, Proteinuria, Protein urine, and Protein urine present
Thyroid Dysfunction	Thyroid dysfunction SMQ broad
Venous Embolic and Thrombotic Events (VTE)	Embolic and thrombotic events, venous SMQ narrow
Palmar-Plantar Erythrodysesthesia (PPE)	MedDRA PTs: Palmar-plantar erythrodysesthesia syndrome, palmar erythema, and plantar erythema.
Fatigue and Asthenic Conditions	MedDRA PTs: Asthenia, Autonomic nervous system imbalance, Decreased activity, Listless, Sluggishness, Chronic fatigue syndrome, Fatigue and Malaise)
Hepatic Disorders	Cholestasis and jaundice of hepatic origin (SMQ narrow) Drug-related hepatic disorders - comprehensive search (SMQ narrow) Liver related investigations, signs and symptoms (SMQ narrow) Hepatic & Hepatobiliary disorders (HLGT)
Congestive Heart Failure/Cardiomyopathy	MedDRA PTs: Acute left ventricular failure, Acute right ventricular failure, Cardiac failure, Cardiac failure acute, Cardiac failure chronic, Cardiac failure congestive, Cardiac failure high output, Cardiogenic shock, Cardiopulmonary failure, Chronic left ventricular failure, Chronic right ventricular failure, Cor pulmonale, Cor pulmonale acute, Cor pulmonale chronic, Ejection fraction decreased, Left ventricular failure, Low cardiac output syndrome, Neonatal cardiac failure, Right ventricular failure, Ventricular failure, Cardiac output decreased, Cardio-respiratory distress, Central venous pressure increased, Diastolic dysfunction, Left ventricular dysfunction, Myocardial depression, Oedema due to cardiac disease, Right ventricular dysfunction, Systolic dysfunction, Ventricular dysfunction, Cardiomyopathy acute, Congestive cardiomyopathy, Ejection fraction abnormal, Ejection fraction decreased.
Carcinogenicity (Second Primary Malignancy)	SMQ malignant or unspecified tumours, SMQ Malignant tumours, SMQ Tumours of unspecified malignancy and the Lower Level Term (LLT): Secondary primary malignancy.

For Cardiovascular disorders: Include Cardiac disorders, Nervous system disorders, Vascular disorders

7.5.2. Clinical Laboratory Evaluation

Laboratory tests to be performed, and timing of collection, are detailed in Section 7.2.2 of the protocol and include the following:

- Hematology: Hgb, white blood cell count (WBC), absolute neutrophil count (ANC), lymphocyte count and platelet count.

- Chemistry: Total bilirubin, aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase, albumin, total protein, lactate dehydrogenase (LDH), sodium, potassium, chloride, calcium, phosphate, blood urea nitrogen (BUN)/Urea, creatinine, glucose.

- Urine dipstick for protein (subjects with ≥2+ protein will have urine protein creatinine [UPC] ratios).

- Pregnancy test, if applicable. Urine pregnancy test will be conducted at sites for females of childbearing potential only. If positive, then serum test should be conducted at central laboratories.

- Thyroid Function: TSH, free T3 and free T4

Baseline laboratory values will be selected from the date closest to but on or prior to the first dose (if there is more than one baseline evaluation).

Summary statistics for actual values and change from baseline will be presented for all planned time points for Hematology, Chemistry, Urinalysis and Thyroid function tests.

Shift from baseline to all post-baseline time points will be produced for urine protein.

Shift from baseline to all post-baseline time points in toxicity grading will be produced for all laboratory parameters which are graded per NCI CTCAE, along with shift from baseline to worst toxicity grade (where worst toxicity grade is defined as the worst toxicity grade experienced at any post-baseline time point, including unscheduled assessments). For laboratory parameters that are graded in both directions e.g. “hyper” and “hypo”, separate rows will be presented for each directional assessment. The hematologic and chemistry laboratory results will be graded according to the NCI CTCAE v4.03 severity grade. For parameters for which an NCI CTCAE v 4.03 scale does not exist, the frequency of subjects with values below, within and above the normal range for the local lab will be summarized.

A summary table of the proportion of subjects meeting Hy’s Law criteria will be provided. The following algorithm will be utilized to determine if Hy’s Law Criteria have been met:

1) if ALT|AST baseline is missing, low or normal: Alkaline Phosphatase < 2*Upper limit of normal (ULN) AND Bilirubin >= 2*ULN AND ALT|AST >= 3*ULN;

2) if ALT|AST baseline is high: Alkaline Phosphatase < 2*ULN AND Bilirubin >= 3*ULN or (Bilirubin >= 2*ULN and Bilirubin >= ULN+Baseline) AND ALT|AST >= 8*ULN or (ALT|AST >= 3*ULN and ALT|AST >= 2*Baseline)

A Hy’s Law Listing will be provided.

All laboratory results will be listed. Any unscheduled laboratory assessments will be listed only.

7.5.3. ECG Evaluations

ECG evaluations are performed at screening only in this study.

The number and percentage of subjects with Normal, Abnormal not Clinically Significant, Abnormal Clinically Significant and Unevaluable reading at screening will be summarized by treatment group.

Summary statistics for QT Interval, Heart Rate, RR Interval, QT corrected for heart rate using Bazett's method (QTcB) and QT corrected for heart rate using Fridericia's method (QTcF) will be produced by treatment group using categories as defined in the International Conference on Harmonisation (ICH) E14 guideline (i.e., QTcF millimeter/second [ms]: ≤ 450; >450 - ≤ 480; >480-≤ 500; >500).

All ECG results will be listed.

7.5.4. Vital Signs, Physical Findings, and Other Observations Related to Safety

Baseline for body weight and vital signs will be defined as the latest value recorded on or prior to dosing on Day 1. Summary statistics for actual values and change from baseline will be presented. Unscheduled body weight and vital signs measurements will be listed only.

The proportion of subjects meeting the following blood pressure criteria on 2 or more scheduled and unscheduled visits (need not be consecutive) after first dose (modified from JNC criteria to address single measurement per time point) (Chobanian et al., 2003)[6] will be presented in a table:

- Normal: Systolic Blood Pressure (SBP) <120 mmHg and Diastolic Blood Pressure (DBP) < 80 mmHg
- Pre-hypertension*: SBP 120-139 mmHg or DBP 80-89 mmHg
- Stage 1: SBP 140-159 mmHg or DBP 90-99 mmHg
- Stage 2: (SBP ≥160 mmHg and DBP <120) or DBP 100-119 mmHg

*Defined according to CTCAE v4.03 criteria.

The proportion of subjects with weight loss ≥10% after first dose will also be presented in a table.

ECOG PS and change from baseline in ECOG PS will be summarized using summary statistics, and will also be presented as a shift from baseline to post-baseline time points. Specifically, shift tables of ECOG score from baseline will be presented at each cycle as well as tables demonstrating from baseline to best and worst post baseline score. A mixed model will also be implemented to assess treatment differences in ECOG PS over time. Baseline ECOG will be

included in the model as a covariate

Physical examinations will be listed only.

7.5.5. Prior/Concomitant Medications

Prior and concomitant medications will be coded using the WHO-drug (World Health Organization) coding dictionary.

Prior medications are defined as those medications stopped prior to the first day of study treatment. Concomitant medications are defined as those medications either stopped, ongoing or started on or after the first day of study treatment up to 28 days post the last day of study treatment.

Number and percentage of subjects taking all prior and concomitant medications will be summarized by ATC Class and WHO Drug Name.

A listing will be provided for prior and concomitant non-drug treatments and procedures.

Prior non-drug treatments and procedures are defined as those stopped prior to the first day of study treatment. Concomitant non-drug treatments and procedures are defined as those treatments and procedures that are either stopped, ongoing or started on or after the first day of study treatment.

Prior treatment for cancer under study (including drugs administered, start and stop date of each, and reason for taking the medication) will be provided in a listing.

Anti-tumor drug treatments taken at follow-up will also be listed and summarized using the WHO-drug coding dictionary. Start dates to be considered as occurring during follow-up will be those that are >28 days beyond the last day of study treatment.

8. References

1. AG-013736 Protocol, Version 10, 21MAR2017.

2. Brookmeyer R and Crowley JJ. A confidence interval for median survival time. Biometrics, 38, 29-41, 1982.

3. Schoenfeld D. Chi-squared goodness of fit tests for the proportional hazards regression model. Biometrika, 1980, 67:145-153.

4. Zhang, Xu. Comparison of restricted mean survival times between treatments based on a stratified Cox model. Bio-Algorithms and Med-Systems 2013; 9(4): 183–189.

5. Amit O, et al. Blinded independent central review of progression in cancer clinical trials: Results from a meta-analysis and recommendation from a PhRMA working group. European Journal of Cancer 47:1772-1778, 2011.

6. Chobanian AV, Bakris GL, Black HR et. al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. JAMA, 2003, 289(19):2560-72.

9. Appendix A

Table 1. Primary Endpoint Disease-Free Survival by Independent Review Committee
Censoring Rules:

		Date of event*/ censoring	**Censoring**
No IRC confirmed recurrence or occurrence of secondary malignancy or death	Post baseline assessments performed	Date of last IRC reviewed scan prior to receiving anti-tumor therapy, if applicable	Yes
	No post baseline assessments performed	Date of randomization	Yes
IRC confirmed recurrence, secondary malignancy or death	Baseline scan shows presence of disease	Date of randomization	No
	Further anti-tumor therapy received after randomization and before recurrence or occurrence of secondary malignancy or death (Subject does not have two or more consecutively missed/not readable IRC reviewed scans immediately prior to recurrence, secondary malignancy or death)**	Date of last IRC reviewed scan prior to receiving anti-tumor therapy	Yes
	Further anti-tumor therapy received after randomization and before recurrence or occurrence of secondary malignancy or death and subject had two or more consecutively missed/not readable IRC reviewed scans immediately prior to recurrence, secondary malignancy or death)**	Date of last IRC reviewed scan prior to the two consecutively missed/not readable scans and prior to receiving anti-tumor therapy**	Yes
	No anti-tumor therapy received after randomization (Subject does not have two or more consecutively missed/not readable IRC reviewed scans immediately prior to recurrence, secondary malignancy or	Date of IRC confirmed recurrence or occurrence of secondary malignancy or death, whichever occurred first*	No

	death)**		
	No anti-tumor therapy received after randomization but subject had two or more consecutively missed/not readable IRC reviewed scans immediately prior to recurrence, secondary malignancy or death**	Date of last IRC reviewed scan before two consecutively missed/not readable IRC reviewed scans **	Yes

* Note: For subjects for whom there is site confirmation of recurrence or occurrence of secondary malignancy in the absence of IRC imaging confirmation, the date of IRC confirmation from histo-/cytopathology specimens will be selected as the date of event if available. For subjects for whom the IRC sees a finding on a scan that requires confirmation but an additional scan is not provided, if there is a subsequent pathology report available the pathology report could be the confirmation of the disease seen on the scan. In this case the scan date should be used as the date of recurrence or secondary malignancy.

**Additional details regarding the definition of two or more consecutively missed or inadequate scans can be found in Appendix B. For subjects who died prior to missing two scheduled assessments, they will be coded as an event at the date of death and will not be censored as long as no anti-tumor therapy was given after randomization and before death.

Table 2. Secondary Endpoint Overall Survival Censoring Rules:

		Date of event/ censoring	**Censoring**
Alive	Post baseline assessments performed	Date subject last known to be alive	Yes
	No post baseline assessments performed	Date of randomization	Yes
Dead		Date of death	No

Table 3. Sensitivity Analysis Disease-Free Survival by Investigator Assessment Censoring Rules:

		Date of event/ censoring	**Censoring**
No Investigator confirmed recurrence or occurrence of secondary malignancy or death	Post baseline assessments performed	Date of last local imaging assessment prior to receiving anti-tumor therapy if applicable	Yes
	No post baseline assessments performed	Date of randomization	Yes
Investigator confirmed	Baseline scan evaluated by the investigator shows	Date of randomization	No

recurrence or secondary malignancy or death	presence of disease		
	Further anti-tumor therapy received after randomization and before recurrence or occurrence of secondary malignancy or death (Subject does not have two or more consecutively missed local imaging assessments immediately prior to recurrence, secondary malignancy or death)*	Date of last local imaging assessment prior to receiving anti-tumor therapy	Yes
	Further anti-tumor therapy received after randomization and before recurrence or occurrence of secondary malignancy or death and subject had two or more consecutively missed local imaging assessments immediately prior to recurrence, secondary malignancy or death)*	Date of last local imaging assessment prior to the two consecutively missed scans and prior to receiving anti-tumor therapy*	Yes
	No anti-tumor therapy received after randomization (Subject did not have two or more consecutively missed local imaging assessments)*	Date of investigator confirmed recurrence or occurrence of secondary malignancy or death, whichever occurred first	No
	No anti-tumor therapy received after randomization but subject had two or more consecutively missed local imaging assessments*	Date of last local imaging assessment before two consecutively missed local imaging assessments*	Yes

*Additional details regarding the definition of two or more consecutively missed scans can be found in Appendix B. For subjects who died prior to missing two scheduled assessments, they will be coded as an event at the date of death and will not be censored as long as no anti-tumor therapy was given after randomization and before death.

Table 4. Sensitivity Analysis Disease-Free Survival by Independent Review Committee without Censoring Subjects for Anti-tumor Therapy or for two or more consecutively missed or not readable scans
Censoring Rules:

		Date of event/ censoring	Censoring
No IRC confirmed recurrence or occurrence of secondary malignancy or death	Post baseline assessments performed	Date of last IRC reviewed scan	Yes
	No post baseline assessments performed	Date of randomization	Yes
IRC confirmed recurrence or occurrence of secondary malignancy or death	Baseline scan shows presence of disease	Date of randomization	No
		Date of IRC confirmed recurrence or occurrence of secondary malignancy or death, whichever occurred first*	No

* Note: For subjects for whom there is site confirmation of recurrence or occurrence of secondary malignancy in the absence of IRC imaging confirmation, the date of IRC confirmation from histo-/cytopathology specimens will be selected as the date of event if available. For subjects for whom the IRC sees a finding on a scan that requires confirmation but an additional scan is not provided, if there is a subsequent pathology report available the pathology report could be the confirmation of the disease seen on the scan. In this case the scan date should be used as the date of recurrence or secondary malignancy.

Table 5. Sensitivity Analysis Disease-Free Survival by Independent Review Committee with Considerations for Start of New Anti-Tumor Therapy as Events
Censoring Rules:

		Date of event*/ censoring	Censoring
No IRC confirmed recurrence or occurrence of secondary malignancy or death	Post baseline assessments performed, subject did not start a new anti-tumor therapy	Date of last IRC reviewed scan.	Yes
	Post baseline assessments performed, subject did start a new anti-tumor therapy	Date of start of anti-tumor therapy.	No
	No post baseline assessments performed	Date of randomization	Yes
IRC confirmed recurrence or secondary	Baseline scan shows presence of disease	Date of randomization	No
	Subject did not start a new	Date of IRC confirmed	No

malignancy or death	anti-tumor therapy (Subject did not have two or more consecutively missed/not readable IRC reviewed scans immediately prior to recurrence, secondary malignancy or death)**	recurrence or occurrence of secondary malignancy or death, whichever occurred first	
	Subject did not start a new anti-tumor therapy (Subject had two or more consecutively missed/not readable IRC reviewed scans immediately prior to recurrence, secondary malignancy or death)**	Date of last IRC reviewed scan before two consecutively missed/not readable IRC reviewed scans**	Yes
	Subject did start a new anti-tumor therapy before IRC confirmed recurrence or secondary malignancy or death (Subject did not have two or more consecutively missed/not readable IRC reviewed scans immediately prior to recurrence, secondary malignancy or death or immediately prior to new anti-tumor therapy**	Date of start of new anti-tumor therapy	No
	Subject did start a new anti-tumor therapy before IRC confirmed recurrence or secondary malignancy or death (Subject had two or more consecutively missed/not readable IRC reviewed scans immediately prior to new anti-tumor therapy)**	Date of last IRC reviewed scan before two consecutively missed/not readable IRC reviewed scans **	Yes

* Note: For subjects for whom there is site confirmation of recurrence or occurrence of secondary malignancy in the absence of IRC imaging confirmation, the date of IRC confirmation from histo-/cytopathology specimens will be selected as the date of event if available. For subjects for whom the IRC sees a finding on a scan that requires confirmation but an additional scan is not provided, if there is a subsequent pathology report available the pathology report could be the confirmation of the disease seen on the scan. In this case the scan date should be used as the date of recurrence or secondary malignancy.

**Additional details regarding the definition of two or more consecutively missed or inadequate scans can be found in Appendix B. For subjects who died prior to missing two scheduled assessments, they will be coded as an event at the date of death and will not be censored.

Table 6. Disease-Free Survival by Independent Review Committee Censoring Only at Scheduled Visits
Censoring Rules:

		Date of event*/ censoring	**Censoring**
No IRC confirmed recurrence or occurrence of secondary malignancy	Post baseline assessments performed	Date of last scheduled IRC reviewed scan prior to receiving anti-tumor if applicable	Yes
	No post baseline assessments performed	Date of randomization	Yes
IRC confirmed recurrence or secondary malignancy or death	Baseline scan shows presence of disease	Date of randomization	No
	Further anti-tumor therapy received after randomization and before recurrence or occurrence of secondary malignancy or death (Subject did not have two or more consecutively missed/not readable IRC reviewed scans)	Date of last scheduled IRC reviewed scan prior to receiving anti-tumor therapy	Yes
	Further anti-tumor therapy received after randomization and before recurrence or occurrence of secondary malignancy or death (Subject did have two or more consecutively missed/not readable IRC reviewed scans)	Date of last scheduled IRC reviewed scan prior to receiving anti-tumor therapy and before two consecutively missed/not readable IRC reviewed scans	Yes
	No anti-tumor therapy received after randomization (Subject did not have two or more consecutively missed/not readable IRC reviewed scans)	Date of scheduled IRC assessment which confirmed recurrence or occurrence of secondary malignancy. If event is death then the date of death will be used.	No
	No anti-tumor therapy received after randomization (Subject did have two or more consecutively missed/not readable IRC reviewed scans)	Date of last scheduled scan before two consecutively missed/not readable IRC reviewed scans	Yes

*Events that occur within +/-4 weeks of a scheduled scan will be considered an event at the scheduled scan. Events outside the 4-week window will be counted as events at the next scheduled scan time and

censoring outside the 4-week window would be censored at the previous scan.

Table 7. Disease-Free Survival by Independent Review Committee Using a Parametric Model for Interval-Censored Data
Censoring Rules:

		Date of event*/ censoring	**Censoring**
No IRC confirmed recurrence or occurrence of secondary malignancy	Post baseline assessments performed	Date of last IRC reviewed scan prior to receiving anti-tumor therapy if applicable	Yes
	No post baseline assessments performed	Date of randomization	Yes
IRC confirmed recurrence or secondary malignancy or death	Baseline scan shows presence of disease	Date of randomization	No
	Further anti-tumor therapy received after randomization and before recurrence or occurrence of secondary malignancy or death (Subject did have two or more consecutively missed/not readable IRC reviewed scans)	Date of midpoint between last IRC reviewed scan prior to receiving anti-tumor therapy and before two consecutively missed/not readable IRC reviewed scans and date of IRC confirmed recurrence or secondary malignancy or death	Yes
	No anti-tumor therapy received after randomization (Subject did have two or more consecutively missed/not readable IRC reviewed scans)	Date of midpoint between last IRC reviewed scan prior to two consecutively missed/not readable scans where there was no IRC confirmed recurrence and date of IRC confirmed recurrence or occurrence of secondary malignancy or death	Yes
	Further anti-tumor therapy received after randomization and before recurrence or occurrence of secondary malignancy or death (Subject did not have two or more consecutively missed/not readable IRC reviewed scans)	Date of midpoint between last IRC reviewed scan prior to receiving anti-tumor therapy and date of IRC confirmed recurrence, occurrence of secondary malignancy or death	Yes
	No anti-tumor therapy received after randomization (Subject did not have two or more consecutively missed/not readable IRC reviewed scans)	Date of IRC confirmed recurrence or occurrence of secondary malignancy or death	No

10. Appendix B

Data Handling for DFS Events That Occurred after Two or More Consecutively Missed or Not Readable Scans

Section 5.1 of the SAP states the following: "For subjects who had two or more consecutively missed or not readable IRC reviewed scans immediately prior to a recurrence or occurrence of a secondary malignancy or death,

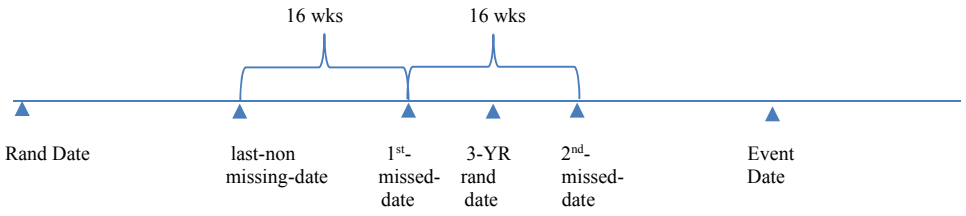
- DFS will be censored on the date of the last IRC reviewed scan prior to the consecutively missed or not readable IRC reviewed scans."

Subjects in this trial are being followed for recurrence or occurrence of a secondary malignancy every 16 weeks during the first 3 years of treatment after randomization and every 6 months after the first 3 years of treatment. Since the frequency of tumor scans varies depending on whether the tumor scan is within 3 years from randomization or after 3 years from randomization, the length of time interval for two consecutively missed tumor scans can be different.

There are 3 scenarios to determine whether a DFS event should be censored due to 2 or more consecutively missed tumor scans. In order to simplify the description of the 3 scenarios, the following abbreviations are defined:

Definition	Abbreviation
date of the last evaluable scan prior to the DFS event	last-non-missing-date
date for the first missed tumor scan that is after the last non-missing tumor scan prior to the DFS event	1st -missed-date
date for the 2nd consecutively missed tumor scan that is after the last non-missing tumor scan prior to the DFS event	2nd - missed date
date of 3 years after randomization date	3-Year (YR) rand date

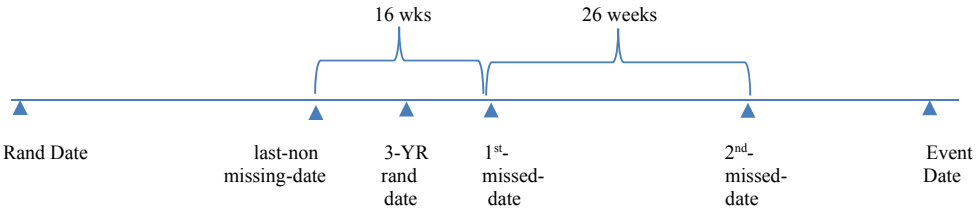
(1) The last-non-missing-date is 16 weeks or more before the 3-YR rand date.



In Scenario (1), the targeted date for the 1st-missed-date is 16 weeks after the last-non-missing-date,

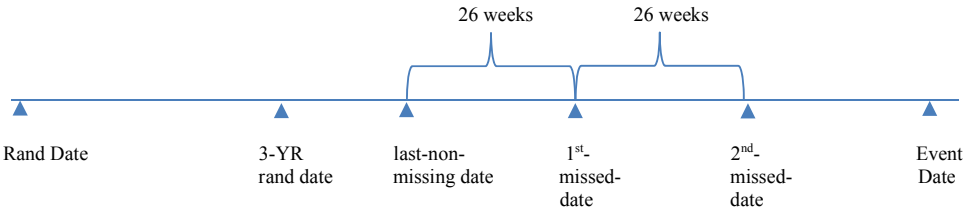
and the targeted date for the 2nd-missed-date is 32 weeks after the last non-missing-date. After adding a 2-week window, the decision rule is: if the DFS event is >34 weeks after the last-non-missing date, DFS time will be censored at the date of the last non-missing tumor scan prior to the DFS event, otherwise DFS time will be the non-censored time interval from randomization date to the date of DFS event.

(2) The last-non-missing-date is before the 3-YR rand date, but the last-non-missing-date plus 16 weeks is beyond the 3-YR rand date.



In Scenario (2), the targeted date for the 1st-missed-date is 16 weeks after the last-non-missing-date, and the targeted date for the 2nd-missed-date is 16 weeks plus 26 weeks after the last-non-missing-date. After adding a 2-week window, the decision rule is: if the DFS event is >44 weeks) after the last-non-missing-date, DFS time will be censored at the date of the last non-missing tumor scan prior to the DFS event, otherwise DFS time will be the non-censored time interval from the randomization date to the date of DFS event.

(3) The last-non-missing-date is beyond the 3-YR rand date.



In Scenario (3), the targeted date for the 1st- missed-date is 26 weeks after the last-non-missing-date, and the targeted date for the 2nd-missed-date is 52 weeks after the last-non-missing-date. After adding a 4-week window, the decision rule is: if the DFS event is >56 weeks after the last non-missing-date, DFS time will be censored at the date of the last non-missing tumor scan prior to the DFS event, otherwise DFS time will be the non-censored time interval from randomization to the date of DFS.

If no DFS event occurred after the 2 or more consecutively missed tumor scans, DFS time will be the censored time interval from randomization date to the last tumor scan date regardless whether these scans were done after 2 or more consecutively missed tumor scans.

11. Table Shells and Specifications

11.1. Table Shells

Table shells are provided in a separate document.

12. Figure Shells and Specifications

12.1. Figure Shells

Listing shells are provided in a separate document.

13. Listing Shells and Specifications

13.1. Listing Shells

Listing shells are provided in a separate document.

14. Programming Conventions for Outputs:

Partial dates should be presented as - -NOV1999 or - - - - - 1999 as needed.

Partial times should be presented as --:30 or 14:-- as needed.

Listings should be sorted according to the order of the columns.