A Multi-Center, Open Label, Randomized Phase 2 Study of AGS-16C3F vs. Axitinib in Metastatic Renal Cell Carcinoma

ISN/Protocol AGS-16C3F-15-3

ClinicalTrials.gov Identifier: NCT02639182

Date of SAP v2: 10 Jun 2019

Sponsor: Astellas Pharma Global Development, Inc. (APGD)

1 Astellas Way

Northbrook, IL 60062

STATISTICAL ANALYSIS PLAN

PROTOCOL AGS-16C3F-15-3

A MULTI-CENTER, OPEN LABEL, RANDOMIZED PHASE 2 STUDY OF AGS-16C3F VS. AXITINIB IN METASTATIC RENAL CELL CARCINOMA

AUTHOR: PPD

VERSION NUMBER AND DATE: V2.0 FINAL, 10 JUNE 2019

STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

Statistical Analysis Plan V2.0 (Dated 10 June 2019) for Protocol AGS-16C3F-15-3.

	Name	Signature	Date (DD/MM/YYYY)
Author:	PPD	PPD	
Position:	PPD		
Company:	Quintiles		
Reviewer:	PPD		
Position:	PPD		
Company:	Astellas Pharma Global Development		
Reviewer:	PPD		
Position:	PPD		
Company:	Astellas Pharma Global Development		

Reviewer:	PPD		
Position:	PPD		
Company:	Astellas Pharma Global Development		
Reviewer:	PPD		
Position:	PPD		
Company:	Astellas Pharma Global Development		
Reviewer:	PPD		
Position:	PPD		
Company:	Astellas Pharma Global Development		
Reviewer:	PPD		

Position:	PPD		
Company:	Astellas Pharma Global Development		
Reviewer:	PPD		
Position:	PPD		
Company:	Astellas Pharma Global Development		
Reviewer:	PPD		
Position:	PPD		
Company:	Astellas Pharma Global Development		
Reviewer:	PPD		
Position:	PPD		
Company:	Astellas Pharma Global Development		
Reviewer:	PPD		
Position:	PPD		
Company:	Astellas Pharma Global Development		

Upon review of this document, the undersigned approves this version of the Statistical Analysis Plan (Dated 10 Jan 2018), authorizing that the content is acceptable for the reporting of this study.

	Name	Signature	Date (DDMMMYYYY)
Approved By:	PPD		
Position:	PPD		
Company:	Astellas Pharma Global Development		
Approved By:	PPD .		
Position:	PPD		
Company:	Astellas Pharma Global Deveopment		
Approved By:	PPD		
Position:	PPD		
Company:	Astellas Pharma Global Development		

MODIFICATION HISTORY

Unique Identifier for this Version	Date of the Document Version	Author	Significant Changes from Previous Version
0.1	15JAN2016	PPD	SAP Draft 1 based on Original Protocol (dated 02 Jan 2016).
0.2	25JUL2016	PPD	SAP Draft 2 updated to match Amendment 1 (dated 28 Jan 2016) and Amendment 2 (27 May 2016).
0.3	13May2017	PPD	SAP Draft 3 incorporates sponsor comments on Draft 2 through 13 May 2017.
0.4	06 Oct 2017	PPD	SAP Draft 4 updated to match Amendment 3 (dated 21 Aug 2017). Incorporates sponsor comments on Draft 3 through 26 Sep 2017.
0.5	27 Oct 2017	PPD	Incorporates sponsor comments on Draft 4 through 08 Nov 2017.
0.6	08 Dec 2017	PPD	Incorporates sponsor comments on Draft 5 through 01 Dec 2017.
0.7	14 Dec 2017	PPD	Incorporates sponsor comments on Draft 6 through 14 Dec 2017. Final document clean-up.
0.8	18 Dec 2017	PPD	Fix typo in ECG section. Final document clean-up.
V1.0 Final	10 Jan 2018	PPD	Incorporate sponsor comments on Draft 8 through 10 Jan 2018. Final document clean-up.
V2.0 final	10 June 2019	PPD	Incorate comments from Sponsor

TABLE OF CONTENTS

1. INTRODUCTION

This document describes the rules and conventions to be used in the presentation and analysis of efficacy, safety, pharmacokinetic, and antibody formation data for Protocol AGS-16C3F-15-3. It describes the data to be summarized and analyzed, including specifics of the statistical analyses to be performed.

PK parameter estimation is outside the scope of this statistical analysis plan (SAP) and may be addressed by the sponsor in a separate document. Reporting of the PK concentrations and estimated parameters will be addressed in this SAP.

This SAP is based on Protocol AGS-16C3F-15-3, version 4.0 (Amendment 3, dated 18 Jul 2017) and AGS-16C3F-15-3 Subject Case Report Forms, Version 6.0 (dated 26 SEP 2017).

2. STUDY OBJECTIVES

2.1. PRIMARY OBJECTIVE

To evaluate the Progression Free Survival (PFS), based on investigator radiologic review, of AGS-16C3F compared to axitinib in subjects with metastatic renal cell carcinoma.

2.2. SECONDARY OBJECTIVES

- To evaluate the following for AGS-16C3F compared against axitinib:
  - PFS per RECIST v1.1 by blinded central radiology assessment
  - Overall Survival (OS)
  - Objective Response Rate (ORR) based on the investigator's radiographic assessment
  - Disease Control Rate (DCR) based on the investigator's radiographic assessment
  - Duration of Response (DOR) based on the investigator's radiographic assessment
  - Safety
- To evaluate the pharmacokinetics of AGS-16C3F
- To evaluate the incidence of antidrug antibody formation to human native antibody (AGS-16C3) and antibody drug conjugate (AGS-16C3F)

2.3. SAFETY OBJECTIVES

- To characterize the general safety profile of AGS-16C3F compared to axitinib, based on adverse event reports, vital sign measurements, electrocardiogram (ECG) results, and clinical laboratory results
- To monitor ocular safety, based on complete eye examinations, ocular symptoms, and clinical ocular adverse events

3. STUDY DESIGN

3.1. OVERVIEW

This multi-center, open-label, randomized study will enroll subjects with metastatic RCC of all histologies who have evidence of progression on or after the last regimen received and have also received at least 2 prior systemic regimens, 1 of which is an anti-VEGF agent (clear cell subjects). The study duration, up to the study's primary endpoint assessment, is approximately 24 months.

Non-clear cell RCC is a heterogeneous group of tumors typically not driven by von Hippel-Lindau (VHL) gene and sensitivity to anti-VEGF treatments differ from that shown in clear cell RCC. Hence, subjects with non-clear cell histology are only required to have progressed on or after only at least one anti-VEGF regimen. Subjects with non-clear cell histology must have ENPP3 positive immunohistochemical (IHC) staining at pre-screening. Positivity is defined as an IHC H-score ≥15. Non-clear cell subjects whose tissue has an IHC H-score of ≥15 will qualify to screen for the study. Clear cell subjects are not required to be screened for ENPP3 to qualify for the study, but subjects must submit tissue during the study for retrospective target expression assessment by IHC.

Participants in this clinical investigation shall be referred to as "subjects". This study will enroll approximately 134 subjects. Within this total (i.e., 134), the number of subjects with non-clear cell histology is limited to 26 total.

Subjects will be randomized at 1:1 ratio to either AGS-16C3F or axitinib. Randomization will be stratified according to ECOG Performance Status (0 or 1), the number of prior systemic RCC regimens (2 or >2), and RCC histology (clear cell or non-clear cell). For stratification purpose:

- Non-clear cell subjects who have progressed after only one prior anti-VEGF regimen will be regarded the same as those clear cell subjects who have progressed after 2 prior systemic RCC regimens.
- Non-clear cell subjects who have progressed after >1 prior systemic RCC regimens will be regarded the same as those clear cell subjects who have progressed after >2 prior systemic RCC regimens.

AGS-16C3F will be administered at 1.8 mg/kg as a single 60-minute IV infusion once every 3 weeks. The dosage in mg will be calculated using the subject's actual body weight in kilograms at predose Cycle 1 Day 1 (C1D1). The same dose will be used throughout the study unless the subject's weight changes from baseline by ≥10%. In this case, the dose of AGE-16C3F will be administered based on the changed weight and this will become the new baseline weight for the purpose of planned dose recalculation and any further dosing decisions.

Therapy may be interrupted and/or the dose adjusted in case of toxicity. These subjects must also use 1% prednisolone acetate ophthalmic suspension (steroid eye drops) as prophylaxis, starting from Day -1 of each treatment cycle and extending through 7 days after the end of the cycle.

Axitinib is a commercial product and will be administered as defined in its product label. Axitinib will have a starting dose of 5 mg, twice daily continuously, by mouth. Axitinib dose may be adjusted during the study as needed as defined in the product label and local institutional guidelines, in doses ranging from 2 to 10 mg, twice daily continuously, by mouth.

Subjects will be eligible to continue receiving treatment on study until disease progression per RECIST v.1.1, unacceptable toxicity, investigator's decision, or study withdrawal. Subjects will also be discontinued from treatment if a significant non-compliance occurs (e.g., those that compromise study objectives or subject safety), or if a subject is lost to follow-up. The study may also be terminated by the Sponsor for any reason. Refer to Protocol Section 6 for details. All subjects will be required to have a Safety Follow-Up visit at least 28 days (+7 days) after the last dose.

Disease assessments will be performed every 8 weeks (±7 days) counting from C1D1, without regard to any treatment delays. Bone scans will be performed every 12 weeks (±7 days) counting from C1D1 (only subjects with bone disease at screening).

All subjects will be followed for survival approximately every 8 weeks from the Safety Follow-Up visit. This will be done by clinic visit or telephone, until death or study closure, whichever occurs first.

Those subjects who discontinued the study for reasons other than objective disease progression by RECIST v.1.1 will continue to have disease assessments every 8 weeks (±7 days) counting from C1D1. This will continue until subject has radiologically confirmed progression, initiates a new therapy, or subject dies.

The primary efficacy endpoint is PFS based on investigator assessment utilizing RECIST v.1.1. Secondary efficacy endpoints include PFS per blinded central review per RECIST v.1.1, OS, and DCR, ORR, DOR (each of the 3 based on investigator radiographic review). Safety data and PK are also collected.

The study will have an Independent Data Monitoring Committee (IDMC) overseeing the study. The IDMC will conduct safety reviews at predetermined times during the study to make a recommendation for continuation of recruitment, protocol modification or study discontinuation for safety reasons.

The IDMC will minimally consist of a medically qualified individual, an ophthalmologist and a statistician. A separate IDMC Charter will specify the governance and conduct of the IDMC.

3.2. FLOW CHART AND SCHEDULE OF EVENTS

A flow chart of the study design and a Schedule of Events can be found in Protocol AGS-16C3F-15-3 Version 4.0 (Amendment 3), Section V.

3.3. CHANGES FROM PROTOCOL

Changes from the Protocol are cross-referenced in Table A.

Table A. Log of Changes from Protocol

Protocol Section	SAP Section and Change from Protocol	Rationale
Section 2.3.2, Secondary Objectives	Section 2.2, "Secondary Objectives", and throughout: The order of presentation of secondary endpoints changed from the protocol.	Sponsor request
Not Available	Section 2.3: Safety objectives that were implied but not stated in the protocol have been added.	Clarification
Section 4.4, Assignment and Allocation	Section 3.4: Some rephrasing has been done	Clarification
Section 5.1.1.3, Treatment Hiatus	Global: Treatment "hiatus" terminology has been changed to "treatment delay"	Sponsor request
Section 5.1.4.1, Treatment compliance of AGS-16C3F	Section 16: The overdose information has been removed per sponsor request and the compliance with steroid drops has been moved to Section 19.7 (other safety).	Sponsor request

Protocol Section	SAP Section and Change from Protocol	Rationale
Section 5.3.2, Medical History; Section 5.3.3 Diagnosis of the Target Disease, Severity, and Duration of Disease	Section 12: Ocular History has been added to the list of types of history for consistency with the eCRFs. All history has been reorganized to follow the eCRFs: (1) Disease History, (2) Prior Cancer Treatment History (Prior Cancer Systemic Therapy, Prior Cancer Radiotherapy), (3) Ocular History, (4) Medical History	Clarification
Section 7.2.2, Per Protocol Analysis Set	Section 5.2: Details regarding protocol deviations that will result in subject conclusion from the PPS have been added per sponsor communication.	Clarification
Section 7.3.4, Previous and Concomitant Medications	Section 14: Planned summaries of prior medications will not be produced	Sponsor request
Section 7.4, Analysis of Efficacy	Section 4.3, "Main Analysis" refers to the protocol-specified analyses to be conducted afterprimary database hardlock.	Clarification
Not available	Section 4.4: An End of Study analysis has been added	Sponsor request
Section 7.4.1.1, Primary Analysis	Section 17.1.2.1: Protocol-specified censoring rules are identified with respect to the regulatory recommendations and presented as Censoring Rules (Set 1).	Clarification
Section 7.4.1.2, Secondary Analysis	Section 17.1.4 has been titled "Sensitivity Analyses" to reflect the protocol content.	Clarification
Section 7.4.1.2, Secondary Analysis	Detail has been provided by adding Censoring Rules (Sets 2, 3, and 4) to Section 17.1.2.2, Section 17.1.2.3, and Section 17.1.2.4, respectively.	Clarification
Section 7.4.1.3, Subgroup Analysis	Section 17.4: Subgroup analyses in the PPS will not be done.	Sponsor request
Section 7.4.2, Analysis of Secondary Endpoints	Section 17.2.3.3, Section 17.2.3.4, Section 17.2.3.6: For all CMH analyses, the odds ratio will be reported instead of the relative risk for consistency between the estimates which are reported and the tests for treatment effect and homogeneity which are reported.	Sponsor request
Section 7.9, Handling of Missing Data, Outliers, Visit Windows, and Other Information:	Section 7.3: Text and cross-referenced SAP locations of different types of missing data were added.	Clarification
Not available	Section 13 follows the eCRF for Concomitant Procedures. A summary of concomitant procedures will be produced.	Consistency with collected data

Protocol Section	SAP Section and Change from Protocol	Rationale
Not available	Section 15, "Study Drug Exposure", was added. Summaries of AGS-16C3F exposure and axitinib exposure will be produced.	Regulatory requirement
Not available	Section 16.1, "Compliance with Disease Assessment Schedule", was added.	Recommended in regulatory guidance
Not available	Section 17.2.3.1, "Concordance of Disease Assessments by Independent Reviewers", was added.	Sponsor request.

3.4. ASSIGNMENT AND ALLOCATION TO TREATMENT

Randomization will be performed via Interactive Response Technology (IRT). After the subject has completed screening and been deemed eligible, the site staff will access the study's IRT system for subject randomization ≤5 days before C1D1. The IRT system will not allow randomization to occur if it has been more than 28 days since the main ICF has been signed. Specific procedures for randomization through the IRT are described in the study's IRT manual.

Subjects will be randomized at 1:1 ratio to either AGS-16C3F or axitinib. Randomization will be stratified according to ECOG Performance Status (0 or 1), RCC histology (clear cell or non-clear cell), and the number of prior systemic RCC regimens (2 for clear cell or 1 for non-clear cell, >2 for clear cell or >1 for non-clear cell cell).

The RCC histology categorization is defined as the following:

- Clear cell group:
  - Clear cell
  - Clear cell component
  - Mixed clear cell
- Non-clear cell group:
  - Papillary and everything else (e.g., chromophobe, collecting duct, renal medullary, etc., not so common but possible other RCC histologies)

Table B. Categories for Stratification of Prior Systemic RCC for Randomization Process

Prior Systemic RCC Regimen Stratum	Clear Cell Subjects	Non-Clear Cell Subjects
2	2 prior systemic RCC regimens	1 prior systemic RCC regimen
>2	>2 prior systemic RCC regimens	>1 prior systemic RCC regimens

Per protocol, subjects are expected to start study treatment ≤5 days after randomization.

3.5. SAMPLE SIZE JUSTIFICATION

For subjects with metastatic RCC who have progressed after at least 2 prior lines of therapy, a median PFS of 4 months is anticipated for the control agent. It is hypothesized that AGS-16C3F will improve median PFS in such subjects by 2 months (i.e., 4 vs 6 months). Under the assumption of exponential distribution of PFS, such an improvement corresponds to a hypothesized hazard ratio of 0.667. If this hypothesized hazard ratio is true, 110 PFS events provide 80% power to achieve a one-sided p-value of 0.1 or smaller. Such a result will be achieved when the observed hazard ratio is 0.783 or smaller; and the observed difference in median PFS is 1.1 months or larger (assuming observed median PFS for the control arm is 4 months). Assuming accrual of approximately 7 subjects per month for 18 months, 110 PFS events are projected 24 months after the date the first subject is randomized. Approximately 134 subjects may be enrolled to accommodate for up to 10% of subjects who may become lost to follow-up.

4. PLANNED ANALYSES

The following analyses will be performed for this study:

- Analyses for Independent Data Monitoring Committee (IDMC) meetings
- Main Analysis
- End of Study Analysis

A delivery date and an appropriate data cut-off date will be determined by the study team at least 8 weeks in advance of each planned analysis. Data used for planned analyses are expected to be as clean as possible prior to analysis and extracted with the agreement of the sponsor. The QuintilesIMS Data Management Interim Database Lock Plan for this study will provide detail.

4.1. INDEPENDENT DATA MONITORING COMMITTEE (IDMC)

The study will have an Independent Data-Monitoring Committee (IDMC) overseeing the study. The IDMC will conduct safety reviews on a periodic basis during the trial and will make a recommendation for continuation of recruitment, protocol modification or study discontinuation for safety reasons.

It is anticipated that the first IDMC data review meetings will be scheduled to occur via teleconference after the first 25 subjects have been enrolled and participated in study for at least 100 days; subsequent reviews will occur every 6 months. The data which are presented for review will be cumulative, and are expected to be as clean as possible prior to each data snapshot.

The outputs for the safety data review will be a subset of the outputs which are planned for the main analysis. A separate DMC SAP will not be written. A list of the outputs to be generated for each safety review meeting will be provided by Quintiles as a separate document.

Results will have limited distribution and will be disseminated only to the IDMC and to the parties within Agensys and Quintiles who are specified in the study Unblinding Plan. This is to avoid raising significant concerns about introducing potential bias into the conduct of the study or into subsequent decisions regarding the conduct of the study.

The full procedures for IDMC data review will be described in a separate IDMC charter.

4.2. MAIN ANALYSIS

The purpose of the Main Analysis is to establish proof of concept for AGS-16C3F in the treatment of RCC, based on efficacy, safety and PK data; and to support preparation of a Clinical Study Report (CSR).

The Main Analysis is planned after approximately 110 PFS events are observed. The data cut-off date for the Main Analysis will be determined based on the primary efficacy endpoint of PFS under the Censoring Rules (Set 1). Data from visits and assessments that occurred prior to or on the cut-off date are expected to be complete and all critical fields are expected to be clean. Data that occurred after the cut-off date yet entered in the EDC prior to data extraction will be excluded from the main analysis.

The planned analyses identified in this SAP will be performed by Quintiles Biostatistics following sponsor authorization of the SAP, sponsor authorization of the analysis sets, and sponsor permission to lock the database. The study Unblinding Plan provides more detail.

PK parameter estimation will be performed by the Agensys pharmacokineticist and is outside the scope of this SAP. However, Quintiles will report the PK concentration results and PK parameter estimates for the Main Analysis as described in this SAP.

4.3. END OF STUDY ANALYSIS

An End of Study Analysis will be performed at the discretion of the sponsor when any one of the following conditions occurs:

- The IDMC recommends stopping the study
- AGS-16C3F has been either approved or rejected by every country in which subjects were enrolled in the study
- Every randomized subject has met one of the following conditions: (1) the subject has discontinued from the study, or (2) the subject has been followed for a minimum of 12 months after discontinuing treatment, or (3) the subject has been followed for 12 additional months after the planned Main Analysis was performed.

The End of Study Analysis is expected to include safety analyses of AEs and efficacy analyses of the PFS and OS.

5. ANALYSIS SETS

The sponsor will make decisions about subject exclusions from analyses and authorize/approve the analysis set membership prior to each reporting event. Membership in all analysis sets will be determined prior to the Main Analysis, based on all data available at that time. Analysis sets will be updated if appropriate at the time of the End of Study Analysis, and any changes between the Main Analysis and the End of Study Analysis sets will be documented.

5.1. FULL ANALYSIS SET [FAS]

In accordance with the intention-to-treat (ITT) principle, the full analysis set (FAS) will consist of all subjects who are randomized to study drug. This will be the primary analysis set for efficacy analyses. The subjects will be grouped for purposes of the analysis according to the treatment assigned by the randomization, regardless of whether the subject receives the correct drug assignment.

5.2. PER PROTOCOL ANALYSIS SET [PPS]

The per-protocol analysis set (PPS) will consist of all subjects in the FAS who did not experience protocol deviations that might affect the efficacy endpoints. Important protocol deviations that might affect efficacy are defined separately for each version of the protocol, as the inclusion/exclusion numbering may change between versions. Specifically, subjects are expected to be excluded from the PPS due to the following reasons:

- PD1 - Did not meet the inclusion/exclusion criteria that are prognostic of the efficacy endpoints; i.e., one or more of:
  - Original Protocol, Amendment 1, Amendment 2, or Amendment 3: Inclusion Criteria 3, 4, 5, 7, 9, 10 or Exclusion Criteria 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 14.
- PD2 Developed withdrawal criteria during the study and was not withdrawn.
- PD3 - Received wrong treatment or incorrect dose
  - Did not receive treatment
  - Not treated as assigned (i.e., received the wrong drug)
- PD4 – Received excluded concomitant treatment
  - Prohibited concomitant medication [see SAP Appendix 3]
  - Prohibited non-drug therapy [see SAP Appendix 3]

5.3. SAFETY ANALYSIS SET [SAF]

The safety analysis set (SAF) will consist of all randomized subjects who receive at least one dose of study drug (AGS-16C3F or axitinib). Subjects will be classified according to actual study treatment received for the first dose. If there is any doubt whether a subject was treated or not, they will be assumed to have been treated for the purposes of analysis.

The SAF will be used for summaries of all safety variables.

5.4. PHARMACOKINETIC ANALYSIS SET [PKAS]

The pharmacokinetic analysis set (PKAS) will consist of subjects who were randomized to receive AGS-16C3F and who have sufficient serum concentration data to facilitate derivation of at least one PK parameter, and for whom the time of dosing on the day of sampling is known. Additional subjects may be excluded from the PKAS at the discretion of the pharmacokineticist. Any formal definitions for exclusion of subjects or time-points from the PKAS will be documented in the Classification Specifications and determined in the Classification Meeting.

5.5. PHARMACODYNAMIC ANALYSIS SET [PDAS]

A pharmacodynamic marker has not been established for this target.

6. GENERAL CONSIDERATIONS

6.1. STUDY PERIODS

This study has several periods that are defined for analysis as follows:

- Screening: Prior to first dose of study medication; may include pre-screening and screening
- Treatment: From the date of first dose of study drug through the date of last dose or date of decision to end treatment.
- Safety Follow-up: The post-treatment follow-up period which extends from the date of last dose + 1, or the date of decision to end treatment + 1, through the later of the Safety Follow-up visit or (the date of last dose/decision to end treatment plus 28 days).
- Survival Follow-up: From the later of the Safety Follow-up visit or (the date of last dose/decision to end treatment plus 28 days) to death or study closure.

6.2. REFERENCE START DATE AND STUDY DAY

The study day (or "relative day") will be calculated from the reference start date, and will be used to show start/ stop day of assessments and events.

The reference start date is the date of the first dose of study drug. Day 1 is the study day of first dose of study drug. The study day is calculated as shown below, and may appear in listings where an assessment date or event date appears.

- If the date of a visit, assessment, or event is on or after the reference date then study day = (date of event – reference date) + 1.
- If the date of a visit, assessment, or event is prior to the reference date then study day = (date of event – reference date).

In the situation where the date of a visit, assessment, or event is partial or missing, the partial date(s) will be displayed but the study day and any corresponding durations will be missing from the listing(s), unless otherwise specified for a particular display.

6.3. BASELINE

On Cycle 1 Day 1 [C1D1], all procedures other than post-infusion specific procedures must be performed before the first dose of study drug is administered. Subjects must meet all inclusion/exclusion criteria at C1D1, including the results from all predose procedures occurring on C1D1.

Baseline is defined as the last non-missing measurement taken within 28 days prior to reference start date (including unscheduled assessments), unless otherwise specified. In the case where the last non-missing measurement and the reference start date coincide, that measurement will be considered pre-treatment; but Adverse Events (AEs) and medications commencing on the reference start date will be considered post-baseline.

6.4. DERIVED TIMEPOINTS

For subjects in the AGS-16C3F group, each infusion date starts a new cycle. Cycles are numbered consecutively. For subjects in the axitinib group, clinic visits and most other events are identified by the planned study day.

The following derived visit definitions will apply to tables, figures, and listings unless otherwise specified for a particular analysis:

- Worst Post-Baseline: The worst value collected in the period from C1D1 post-dose through the later of the safety follow-up visit or the last dose date + 28 days (inclusive).
- Safety Follow-Up: Everyone is required to have an actual safety follow up visit and one measurement for the safety f/u. For subjects who don't have a safety F/U visit, report the worst value recorded during the period from the day of last dose or decision to end treatment through the subject's last available assessment.
- Final Visit: The chronologically last available assessment.
- Any Time during Treatment/Safety Follow-up: From C1D1 post-dose through the later of the safety follow-up visit or the last dose date + 28 days (inclusive).

Eye assessments are identified in the raw data only by the exam date. Data from the AGS-16C3F subjects will be assigned to a Cycle based on comparing the exam date to the infusion dates. Data from the axitinib subjects will be assigned to a scheduled assessment or identified as a PRN exam, based on the calculated study day of the exam (see Section 6.6, Visit Windowing Conventions).

For tables that show eye-related safety results by timepoint during the study, data from the AGS-16C3F subjects will be grouped for analysis as follows: Day 43 will include data recorded from post-dose C1D1 through Predose C3D1, inclusive; Day 106 will include data recorded from Post-dose C3D1 through Predose C6D1, inclusive; subsequent intervals will be defined similarly at 3-cycle intervals during treatment. For the axitinib subjects, the analysis periods are based on the scheduled/PRN classifications that have been assigned. For both groups, assessments that are made after the end of treatment (EOT) will be grouped together as "EOT and beyond". Both planned and PRN eye assessment data will be included in all intervals.

It may be necessary to derive other timepoint records for some analyses. Where applicable, these are specified for each analysis that requires them.

6.5. RETESTS, UNSCHEDULED VISITS AND EARLY TERMINATION DATA

In general, data recorded at the nominal visit will be presented for by-visit summaries (if applicable). Unscheduled measurements will not be included in by-visit summaries unless otherwise specified for a particular analysis, but will contribute to the End of Cycle, End of Treatment, or End of Study value, or best/ worst case value where required (e.g. shift table).

In the case of a retest (same visit number assigned), the latest available measurement for that visit will be selected for analyses (if applicable). ECGs will be recorded in triplicate; these are not "retest" values. Handling of the ECG data is addressed in Section 19.4.

Early termination data will be mapped to the next available visit number for by-visit summaries.

Listings will include scheduled, unscheduled, retest and early discontinuation data.

6.6. VISIT WINDOWING CONVENTIONS

Subjects in both treatment arms will have disease assessments every 8 weeks (± 7 days), counting from first dose or first infusion, irrespective of any treatment delay. The Schedule of Assessments (Protocol Section V) provides additional information about the procedures and assessments that are to be completed during each visit or Cycle.

The EDC-specified visit designations, and/or Cycle and Cycle Day designations for AGS-16C3F subjects, will be used for listings and for any summaries by visit or timepoint.

If visit windowing is determined to be required for any analysis, the upper limit of each analysis window will be the protocol-specified upper limit of the "acceptable" window for each assessment. The lower limit of the next consecutive window will be set to the following day. Within each window, the assessment closest to the target day and time will be selected for analysis, unless otherwise specified. Other assessments within the window will be treated as unscheduled or PRN assessments. This strategy will help to keep the results aligned with any event-based presentations that may be produced.

6.7. STATISTICAL TESTS

The primary analysis of the primary endpoint will be conducted using a one-sided test with a 0.1 significance level. All other PFS analyses will be conducted using one-sided tests with a 0.1 significance level also. Secondary endpoints such as OS, ORR, DCR, and DOR will be analyzed using two-sided tests with a significance level of 0.05.

6.8. COMMON CALCULATIONS

For quantitative measurements, change from baseline will be calculated as:

- Change from Baseline = Test Value at Visit X – Baseline Value
- Percent Change from Baseline = 100 x (Test Value at Visit X – Baseline Value)/Baseline Value

6.9. SOFTWARE VERSION

All analyses will be conducted using SAS version 9.4 or higher. PK analysis will be conducted by the sponsor, using Phoenix ™ WinNonlin ® Build 6.2.1.51 (Pharsight Corp., Mountain View, CA).

7. STATISTICAL CONSIDERATIONS

7.1. ADJUSTMENTS FOR COVARIATES AND FACTORS TO BE INCLUDED IN ANALYSES

7.1.1. RANDOMIZATION STRATIFICATION FACTORS

The ECOG PS (0, 1), the RCC histological type (clear cell, non-clear cell), and the number of prior systemic RCC regimens (2 for clear cell or 1 for non-clear cell; >2 for clear cell or >1 for non-clear cell) are stratification factors in the randomization.

Two of these stratification factors will be used in the inferential comparisons of the primary and secondary efficacy endpoints between the treatment groups: ECOG PS (0 or 1) and the number of prior systemic RCC regimens (2 for clear cell or 1 for non-clear cell; >2 for clear cell or >1 for non-clear cell). Histology will not be used as a stratification variable in the inferential analyses due to the limited number of non-clear cell subjects expected to be enrolled into the study.

The stratification factor values used to randomize subjects will be obtained from the IRT and will serve to define the strata for the analyses as follows:

- Stratum 1: ECOG-PS = 0 and number of prior systemic RCC regimens = 2 for clear cell or 1 for non-clear cell
- Stratum 2: ECOG-PS = 0 and number of prior systemic RCC regimens = >2 for clear cell or >1 for non-clear cell
- Stratum 3: ECOG-PS = 1 and number of prior systemic RCC regimens = 2 for clear cell or 1 for non-clear cell
- Stratum 4: ECOG-PS = 1 and number of prior systemic RCC regimens = >2 for clear cell or >1 for non-clear cell

7.1.2. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

The following demographic and other baseline characteristics will be used as covariates in exploratory subgroup analyses (Section 7.6):

- Age group
- Gender
- Race
- ECOG-PS at baseline
- RCC Histology
- Number of prior systemic cancer treatment regimens for RCC

7.1.3. PROGNOSTIC RISK GROUP

A prognostic risk group based on the International Metastatic RCC Database Consortium (IMDC) using Heng's 6 prognostic criteria (Heng et al, 2009) is planned for use in selected sensitivity and exploratory subgroup analyses. Heng's prognostic criteria are:

- Karnofsky Performance Score (KPS) <80
- Hemoglobin < lower limit of normal (LLN)
- Time from diagnosis to first line treatment <1 year
- Corrected calcium > upper limit of normal (ULN)
- Platelets > ULN.
- Neutrophil count > ULN

The prognostic risk groups are defined as favorable (0 factors), intermediate (1-2 factors), or poor (≥3 factors).

As the current study collects the ECOG Performance Status rather than the Karnofsky Performance Status (KPS), it is necessary to equate the scores from the two scales. For the purpose of defining the prognostic risk group, the two scales will be mapped as published by Oken et al (1982) and shown in Table C. Accordingly, a Karnofsky score <80 will be considered equivalent to an ECOG score >1.

Table C. Mapping of Karnofsky Performance Score to ECOG Performance Status

Karnofsky Performance Status	ECOG Performance Status
100 - Normal; no complaints	0 – Asymptomatic (Fully active, able to carry on all pre-disease activities without restriction)
90 - Able to carry on normal activity. Minor signs or symptoms of disease 80 - Normal activity with effort	1 – Symptomatic but completely ambulatory (Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light housework, office work)
70 - Care for self. Unable to carry on normal activity or to do active work 60 - Requires occasional assistance, but is able to care for most of his needs	2 – Symptomatic, <50% in bed during the day (Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours)
50 - Requires considerable assistance and frequent medical care 40 - Disabled; requires special care and assistance	3 – Symptomatic, >50% in bed, but not bedbound (Capable of only limited self-care, confined to bed or chair more than 50% of waking hours)
30 - Severely disabled. Hospitalization indicated though death not imminent 20 - Very sick; hospitalization necessary. Active supportive treatment necessary 10 - Moribund	4 – Bedbound (Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair)
0 - Dead	5 - Dead

No other covariates are planned for inclusion in primary, secondary, or sensitivity analyses. However, see Section 7.6 for a more complete list of factors that may be used for exploratory subgroup analyses.

7.2. MULTICENTER STUDIES

Approximately 30 sites (United States, Canada) are planned to participate in this study. Each site is expected to enroll at least 4 subjects in this study. Sites that do not enroll a subject ≤3 months from the study initiation visit date may be terminated and replaced.

Randomization to treatment arms is centralized across all countries and study sites. No examination of treatment by center interaction will be performed.

7.3. MISSING DATA

Table D describes where to find the discussions of missing data imputation rules for different types of data.

Table D. Location of Imputation Rules in This Document

Imputation Type	Notes/Location in this document
Incomplete start or end dates for prior/concomitant medications	Appendix 2
Missing study drug data	Section 15 and Section 16
Missing efficacy data	Section 17, imputation of individual endpoints, including consideration of multiple censoring criteria for use in the analysis of PFS
Incomplete AE data	Section 19.1, missing severity or other event characteristics. Appendix 2, imputation of missing/incomplete dates.
Other safety data	Respective part of Section 19.

7.4. MULTIPLE COMPARISONS/MULTIPLICITY

Analyses will be performed twice during this study. The main analysis will be conducted and conclusions about efficacy will be made when approximately 110 PFS events have occurred. No adjustment of the alpha used in the Main Analysis will be made due to conducting the End of Study Analysis.

As this is a two-arm study with one primary end point, no adjustment for multiple endpoints or multiple comparisons will be required for the primary analysis.

Since the secondary endpoints do not have a natural hierarchy and there is limited power for these tests, nominal p-values will be reported, which means that no formal adjustment for multiplicity will be done. However, if substantial differences in results among these endpoints are observed, reporting of results will include specific statements acknowledging the increase in type 1 error inherent in situations with multiple comparisons.

7.5. ACTIVE-CONTROL STUDIES INTENDED TO SHOW NON-INFERIORITY OR EQUIVALENCE

Not applicable. This is an active-controlled trial designed to show that AGS-16C3F is superior to axitinib. The primary analysis of PFS is based on one-sided test conducted at the .10 level. Details of the primary analysis are provided in Section 17.1.3.

7.6. EXAMINATION OF SUBGROUPS

7.6.1. SUBGROUPS AT BASELINE

Demographics and other baseline characteristics, disease history, prior cancer treatment history, prior cancer systemic therapy, and prior cancer radiotherapy will be summarized by treatment arm and histological group (clear cell, non-clear cell) in the FAS.

7.6.2. SUBGROUPS FOR EFFICACY

This study was not designed to detect treatment differences within subgroups with high statistical power. However, subgroup analyses may be performed for selected baseline characteristics for the purpose of hypothesis generation.

In particular, exploratory subgroup analyses may be conducted to assess potential heterogeneity of treatment effects across levels of baseline characteristics such as age, gender, race, ECOG PS, the number of prior systemic RCC regimens, histological group, and risk group. Analyses will be performed for the FAS. Each analysis will include only treatment, a single exploratory variable, and their interaction. The following subgroups will be assessed.

- Gender
  - Female
  - Male (*)
- Age (years)
  - <65 (*)
  - ≥65
- Race in 2 categories
  - Caucasian (*)
  - Other
- ECOG Performance Status
  - 0
  - 1 (*)
- Histology
  - Clear cell (*)
  - Non-clear cell
- Number of prior systemic RCC regimens
  - 2 if clear cell or 1 if non-clear cell (*)
  - >2 if clear cell or >1 if non-clear cell
- Prognostic risk group (Heng's criteria, see Section 7.1.3)
  - Favorable (0 risk factors)
  - Intermediate (1-2 risk factors)
  - Poor (≥3 risk factors) (*)

Note: (*) Indicates the reference group for statistical models.

8. OUTPUT PRESENTATIONS

Appendix 1 describes the conventions for presentation of data in outputs.

The templates provided with this SAP describe the presentations for this study and therefore the format and content of the summary tables, figures and listings to be provided by Quintiles Biostatistics.

All collected data will be displayed in listings, unless otherwise specified in Section 22.

9. DISPOSITION AND WITHDRAWALS

There is expected to be a pre-screening informed consent and a main informed consent. Only subjects who sign the main consent and are randomized are considered to be enrolled. The set of randomized subjects will be identical to the FAS.

The number and percentage of FAS subjects who are continuing in the study or were discontinued from the study (by primary reason for discontinuation), and the number and percentage who are continuing treatment or were discontinued from treatment (by primary reason for discontinuation)) will be presented by treatment group and overall.

10. PROTOCOL DEVIATIONS AND SUBJECT EXCLUSIONS FROM ANALYSES

All protocol deviations are recorded in the Clinical Trial Management System (CTMS). Only the protocol deviations that are determined to be statistically important protocol deviations, which are defined to include those described in Protocol Section 7.7 and Protocol Section 8.1.6, will be presented on tables and listings for the clinical study report. The statistically important deviations will be summarized by treatment arm and overall, using the number and percentage of subjects who had each type of deviation. A summary by study site will also be produced.

Subject exclusions and reasons for exclusion from analysis sets will be summarized by treatment arm and overall. See also Section 5, "Analysis Sets", for the list of statistically important deviations that could affect efficacy outcomes, as these are the particular important deviations which could result in subject exclusion from some analyses.

11. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Detailed information on subject demographics and other baseline characteristics will be reported. Descriptive statistics will be generated by treatment group and overall. In the FAS, summary statistics will be produced for all FAS subjects and by grouped histological type. In the PPS and SAF, summary statistics will be produced only for all subjects in the analysis set. If the analysis sets are identical the tables will only be presented only once in the CSR. No statistical testing will be carried out for demographic or other baseline characteristics.

The following demographic and other baseline characteristics will be reported for this study:

- Sex (male, female)
- If female, child-bearing potential
- Ethnicity (Hispanic or Latino, not Hispanic or Latino)
- Race (white, black or Afro-American, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, Other)
- Age (years) at Cycle 1 Day 1
- Age group (<65 years, ≥65 years)
- Baseline height (cm)
- Baseline weight (kg)
- BMI (kg/m^2)
- BMI group (<18.5, 18.5 to <30, ≥30)

The following additional baseline characteristics will be reported on the same table:

- ECOG Performance Status (0, 1)
- Grouped histological type (clear cell, non-clear cell)
- Number of prior systemic RCC regimens (2 if clear cell or 1 if non-clear cell, >2 if clear cell or >1 if non-clear cell)
- Prognosis risk group (favorable, intermediate, poor), based on Heng's criteria; see Section 7.1.3.

11.1. Derivations Related to Demographic and Baseline Characteristics

- BMI (kg/m^2) = weight (kg)/height $(m)^2$

12. Surgical and Medical History

Surgical and medical history conditions are defined as those conditions which start prior to the date of first dose of study drug and stop prior to or on the date of first dose of study drug. A detailed surgical and medical history will be obtained for each subject during the screening period. Where applicable, medical conditions and events will be coded to a preferred term using MedDRA version 18.1.

The following types of surgical and medical history will be collected in the eCRFs and reported for the FAS, PPS, and SAF in this study.

- Disease History
- Prior Cancer Treatment History
- Prior Cancer Systemic History
- Prior Cancer Radiotherapy
- Ocular Symptom History
- Medical History

12.1. Disease History

Data collected on renal cell carcinoma diagnosis, metastasis, histology, and additional relevant biological information will be summarized by treatment group and overall. In the FAS, summary statistics will be produced for all FAS subjects and by grouped histological type. In the PPS and SAF, summary statistics will be produced only for all subjects in the analysis set. Each subject's complete cancer history will be listed.

The following information related to renal cell carcinoma diagnosis, metastasis, histology, and additional relevant biological information will be collected on the disease history eCRF.

- Time since diagnosis (years), calculated relative to date of randomization
- Type of solid tumor cancer
- Histological type

- TNM stage at initial diagnosis, based on AJCC 7th edition (Stage I, II, III, IV, Unknown; and individual T, N, M codes)
- Time in years since diagnosis of metastatic disease, calculated relative to date of randomization

Summary statistics will be generated by treatment group and overall. In the FAS, summary statistics will be produced for all FAS subjects and by grouped histological type. In the PPS and SAF, summary statistics will be produced only for all subjects in the analysis set.

Time since diagnosis of RCC and time since diagnosis of metastatic disease will be summarized by treatment arm using descriptive statistics for continuous measures. For other items, the number and percent of subjects who gave each pre-printed response will be presented by treatment arm (see eCRF or mock shells for the list of response choices).

12.2. PRIOR CANCER TREATMENT HISTORY

The following information will be collected on the eCRFs for prior cancer surgery and prior cancer radiotherapy.

- Prior Cancer Surgery
  - Surgery for malignancy prior to enrollment (Y, N)
  - Type of Surgery (Nephrectomy, Partial Nephrectomy, Other (specify))
- Prior Cancer Radiotherapy
  - Radiotherapy for malignancy prior to enrolment (Y, N)
  - Type of Radiation
  - Site
  - Start Date
  - Stop Date

Summary statistics will be generated by treatment group and overall. In the FAS, the summary statistics will be produced for all FAS subjects and by grouped histological type. In the PPS and SAF, the summary statistics will be produced only for all subjects in the analysis set.

The number and percentage of subjects in each treatment arm who had any prior cancer surgery, and the number and percentage who had each type of surgery, will be reported using the pre-printed terms on the eCRF. All data, including the terms listed in the 'other (specify)' category, will be presented in a listing.

The number and percentage of subjects in each treatment arm who had any prior cancer radiotherapy, and the number and percentage who had each type of radiation, will be reported using the pre-printed terms on the eCRF. All data, including the terms listed in the 'other (specify)' category, the radiation site, and the therapy start and stop dates, will be presented in a listing.

12.3. PRIOR CANCER SYSTEMIC THERAPY

The following information will be collected on a Prior Cancer Systemic Therapy eCRF.

- Regimen number
- Treatment name (see eCRF list)
- Start date
- Stop date

Summary statistics will be generated by treatment group and overall. In the FAS, summary statistics will be produced for all subjects in the analysis set and by grouped histological type; this version of the table will also be produced in the SAF for the IDMC. In the PPS and SAF for the clinical study report, summary statistics will be produced only for all subjects in the analysis set.

The number and percentage of subjects who had each actual count of prior systemic RCC regimens (i.e., the number and percentage who had exactly 1 prior regimen, exactly 2 prior regimens, etc.) will be presented. The number and percentage of subjects who had each individual treatment and the number who had each distinct combination of drugs reported as a single regimen will be presented. All data, including the start and stop dates, will be presented in a listing.

12.4. Ocular History

The following data will be collected on an Ocular History eCRF and coded using MedDRA Version 18.1.

- Any ocular history (Y, N)
- Diagnoses
- Start date
- Stop date/ongoing

Summary statistics will be generated by treatment group and overall for the FAS, PPS and SAF.

The number and percentage of subjects who had any ocular history and who had each type of ocular history will be summarized by treatment arm and PT. The PTs will be ordered alphabetically. The diagnoses will be presented as all diagnoses and diagnoses which were ongoing at the time the data were recorded. A data listing will be provided.

12.5. Medical History

Medical history will be captured on the Medical History page of the eCRF and coded using MedDRA Version 18.1. Summary statistics will be produced for the FAS, PPS, and SAF.

The number and percentage of subjects who report any medical history (i.e., “any findings”), any condition in a specified SOC, or who report each unique PT, will be summarized by treatment arm. The SOC’s will be arranged alphabetically. Within SOC, the PTs will be arranged in order of descending frequency in the AGS-16C3F arm. The history will be presented as all medical history and conditions which were ongoing at the time the history was collected. A data listing will be provided.

13. Concomitant Procedures

Concomitant procedures are captured on a separate eCRF. They will be coded and summarized using the same dictionary (MedDRA Version 18.1) and methods as the medical history, as well as presented in a listing.

14. Previous and Concomitant Medications

The protocol indicates that medications including dietary supplements taken from the time of informed consent through the last formal follow-up observational period should be recorded on the Concomitant Medications CRF. All recorded medications will be coded using the WHO drug dictionary version 01 Sep 2015.

Subsequent discussions indicated no interest in medications taken before the first dose of study medication. Therefore, recorded medications that were used on or after C1D1 will be classified as concomitant or post-treatment according to the definitions below; prior medications will be listed but will not be summarized.

Concomitant medications are medications which were used during the period of treatment with study drug, including:

- Medications that started prior to C1D1 and ended (or were ongoing) on or after C1D1
- Medications that started and ended on C1D1
- Medications that started on or after C1D1 and before a decision to discontinue treatment for AGS-16C3F subjects or before the last dose of study medication for axitinib subjects.

Post-treatment (or "post") medications are medications that began on or after the date of a decision to discontinue study drug for AGS-16C3F subjects or after the last dose of study drug for axitinib subjects.

See Appendix 2 for handling of partial dates for medications. The status of each medication as a concomitant or post-treatment medication will be determined after the imputation of any partial or missing dates.

In cases where it is not possible to determine if a medication was used during or after the treatment period, the medication will be classified as a concomitant medication only.

The number and percent of subjects who used each type of concomitant medication, categorized by ATC Level 1 (Anatomical Main Group), ATC Level 4 (Chemical/therapeutic/pharmacological subgroup), and medicinal product (MP), will be summarized by treatment arm for the FAS, PPS and SAF. Post-treatment ('Post') medications taken as part of a subsequent cancer treatment regimen for RCC will be summarized similarly and will be presented on a separate table.

All recorded medication data will be listed.

15. STUDY DRUG EXPOSURE

Algorithms described in this section could change to ensure the correct handling of unusual cases. Any such changes will be documented in the analysis dataset specifications.

All study drug exposure data will be presented in a listing.

15.1. AGS-16C3F

The following study drug parameters will be summarized for subjects who receive AGS-16C3F.

- Duration of treatment (weeks) in the study will be reported using descriptive statistics
- Amount of exposure to study drug during the study, expressed as number of doses received and total amount of drug used (mg), will be reported using descriptive statistics
- Number and percentage of subjects who required a dose reduction to 1.2 mg/kg
- Number and percentage of subjects who had any treatment delay (time since previous dose >28 days) due to toxicity (adverse event)), overall and by Cycle
- Descriptive statistics for the duration of treatment delays (defined as the time between two consecutive dosing visits in days minus 28 days)

15.1.1. DERIVATIONS RELATED TO AGS-16C3F EXPOSURE

15.1.1.1. Dates of First Dose and Last Dose of AGS-16C3F

For subjects who receive AGS-16C3F, the date of first administration of study drug will be the earliest administration date recorded on the AGS-16C3F Drug Administration eCRF. The date of last administration of study drug will be the latest administration date recorded on the AGS-16C3F Drug Administration eCRF.

15.1.1.2. Duration of Treatment with AGS-16C3F

For subjects who are on treatment, the duration of treatment (days) will be calculated as the data cut-off date minus date of the first dose + 1. For subjects who have gone off treatment, the duration of treatment (days) will be calculated as the date decision made to end treatment (from the "Subject Status – Completion" eCRF) minus the date of the first dose + 1. In each case, the calculation will be made without consideration of interruptions, compliance, or dose changes. Delays between doses will be included, unless the subject discontinues from the study without resuming treatment.

Duration of treatment in days will be converted to duration of treatment in weeks, rounded to the closest 1/10) for summary tables and listings, using the following formula: Number of weeks = round ((number of days/7), 0.1).

15.1.1.3. Planned dose of AGS-16C3F

Each subject's initial planned dose (mg) will be calculated as the prescribed dose level (mg/kg) times the baseline body weight (kg). If the subject's weight changes from baseline by ±10% or more during the study, then the new body weight will become the new "baseline" for recalculation of the dose at that visit and for any subsequent calculations of a ±10% change in weight. If the "baseline" body weight or the prescribed dose level changes at any time during treatment, then the planned dose will be recalculated.

15.1.1.4. Amount Taken in Milligrams (AGS-16C3F)

Amount of exposure to study drug will be reported as number of doses (infusions) started and total amount of drug used (mg) per subject. One dose per Cycle is planned.

The need for dose adjustment is collected on the eCRF and will not be imputed if missing.

15.1.1.5. Definition and Calculation of Treatment Delay

The protocol-specified dosing schedule is one dose every 21 days. If for logistical reasons (e.g., public holidays, clinic scheduling, etc.) the next dose cannot be given 21 days from the last dose, a -2 day and a +7 day dosing window is allowed.

Therefore, a delay is defined to have occurred when a subject has >28 days:

- Between 2 consecutive doses, or
- Between the date of the last dose administered and the earlier of (the date of the decision to stop treatment or the data cut-off date).

When a treatment delay occurs, it will be attributed to the later of the two doses. The length of the delay will be calculated as the difference between the two dates, minus the planned cycle time.

For example, suppose that the dates of the Cycle 4 and Cycle 5 doses were 35 days apart. The delay which occurred between Cycle 4 and Cycle 5 would be identified as belonging to Cycle 5 in the analysis dataset. The length of the

delay would be calculated as the date of the Cycle 5 dose minus the date of the Cycle 4 dose minus 28 days; given that the two doses were 35 days apart, this means the length of the delay would be 7 days.

15.2. AXITINIB

Descriptive statistics for the following study drug parameters will be presented for subjects who receive axitinib.

- Duration of treatment per subject, in weeks
- Total amount of medication used per subject, in milligrams
- A flag which indicates that a subject's initial prescribed Axitinib dose of 5 mg was increased at any time during the treatment period.
- A flag which indicates that a subject's initial prescribed Axitinib dose of 5 mg was decreased at any time during treatment.

15.2.1. DERIVATIONS RELATED TO AXITINIB EXPOSURE

15.2.1.1. Dates of First Dose and Last Dose of Axitinib

For subjects who receive axitinib, the date of first dose of axitinib will be the earliest dosing date recorded on the Axitinib Drug Administration eCRF.

The date of last dose will be taken from the "Subject Status – Completion" form. For any subject for whom a last dose date is unavailable at the time of database lock or data cut-off, a last dose date will be imputed as the earliest of:

- The date of discontinuation from study
- The date on which the last dose from the last bottle could have been taken, if the medication had been taken as prescribed; if the imputation of last dose date is based on the last bottle dispensed, then the earlier of the bottle return date or a date based on the total number of doses in the bottle will be used
- The subject's death date
- The day prior to the first dose of any non-study cancer regimen

For subjects who are still on treatment at the time of any interim reporting event, including all IDMC data review meetings and the Main Analysis, the date of last dose may be imputed as the data cut-off date for calculations that require a date of last dose.

15.2.1.2. Duration of Treatment with Axitinib

For subjects who are on treatment, the duration of treatment (days) will be calculated as the data cut-off date minus date of the first dose + 1. For subjects who have gone off treatment, the duration of treatment (days) will be calculated as the last dose date minus the date of the first dose + 1. In each case, the calculation will be made without consideration of interruptions, compliance, or dose changes. Treatment delay time will be included, unless the subject discontinues from the study without resuming treatment.

If the subject discontinues from study, is lost to follow-up, or dies, and the last dose date remains unavailable at the time of database lock or data cut-off, then the imputed date of last dose will be used in calculating the duration of treatment. Duration of treatment in days will be converted to duration of treatment in weeks, rounded to the closest 1/10) for summary tables and listings, using the following formula:

Number of weeks = round ((number of days/7), 0.1).

15.2.1.3. Planned Amount in Milligrams (Axitinib)

Axitinib comes in two strengths (5mg and 1mg tablets) and each subject can be prescribed an axitinib dose ranging from 2 mg to 10 mg per dose, to be taken twice daily. The prescribed amount (mg) per dose is collected in the Axitinib Administration eCRF.

During any dosing interval, the planned amount of axitinib (mg) can be calculated as the prescribed amount (mg) per dose x 2 doses per day x the number of days in the dosing interval. The overall amount (mg) of axitinib planned to be taken by a subject can be calculated by summing the number of planned milligrams across all dosing intervals for the subject.

15.2.1.4. Amount Taken in Milligrams (Axitinib)

In any dosing interval, the amount (mg) taken will be calculated as (the number of 1mg tablets dispensed at the previous visit minus the number of 1mg tablets returned at the current visit) plus 5 x (the number of 5mg tablets dispensed at the previous visit minus the number of 5mg tablets returned at the current visit). The total amount taken (mg) during the study will be calculated by summing the amounts taken by a subject in all dosing intervals.

It will be assumed that pills not returned during the course of the study were taken by the subject as planned, up to and including the date of actual or imputed last dose. It will be assumed that no study drug was taken after the imputed date of last dose.

If a subject fails to return a bottle on schedule, but returns it within 7 days of the planned time, the return will be treated as an on-schedule return for data handling purposes. If a bottle is returned more than 7 days late, then it will be assumed that the subject took the drug on a consistent, personal schedule; and therefore the amounts returned and taken will be averaged over the two affected time intervals.

15.2.1.5. Subject Exposure to Axitinib

Amount of exposure to study drug will be reported as amount of drug used (mg).

15.2.1.6. Determination of Dosing Intervals (Axitinib)

Dosing intervals must be defined to support accurate calculation of the amount of drug planned or taken. In general, a dosing interval will start at a visit when study drug is dispensed and end at the next consecutive visit when study drug is returned. The following data-driven cases may also apply:

- If the prescribed dose changes between visits, the date of the prescription change will signal the end of one dosing interval and start of another.
- If study drug is dispensed or returned at an unscheduled visit, that visit will signal the end of one interval and start of the next.

Other rules for handling special cases may be developed during programming and will be documented in the analysis dataset specifications.

16. COMPLIANCE

16.1. COMPLIANCE WITH STUDY DRUG

Descriptive statistics for overall percent compliance by subject will be presented for the axitinib group in the SAF.

16.1.1. DERIVATIONS RELATED TO STUDY DRUG COMPLIANCE

Compliance will be measured using the overall percentage of planned exposure for each subject. Percent of planned exposure will be calculated as 100 x amount taken (mg)/amount planned (mg). Section 15.2.1 presents the derivations of the planned amount (mg) and amount taken (mg).

16.2. COMPLIANCE WITH DISEASE ASSESSMENT SCHEDULE

A disease assessment is considered compliant with the disease assessment schedule if it is performed within the protocol-specified window. A disease assessment that is performed outside the protocol-specified window, or not performed, is considered to be noncompliant with the protocol-specified schedule.

A descriptive analysis of compliance to the disease assessment schedule will be provided by treatment arm. The number and percentage of off-schedule disease assessments will be calculated and listed for each subject. A table showing the number of subjects with 1 or more off-schedule disease assessments, the total number of off-schedule disease assessments during the study, and descriptive statistics for the number and percentage of off-schedule disease assessments per subject, by treatment arm, will be produced.

16.2.1. DERIVATIONS RELATED TO COMPLIANCE WITH DISEASE ASSESSMENT SCHEDULE

For each subject, the planned disease assessment schedule will be compared to the subject's actual disease assessment schedule. A planned disease assessment will be considered off-schedule if no disease assessment was performed within the planned, protocol-specified assessment window. Missing disease assessments will be included in the counts of off-schedule assessments.

17. EFFICACY

Efficacy analyses will be conducted on the FAS and PPS. The FAS will be considered as primary and the PPS will be used as secondary to assess "sensitivity to analysis sets", i.e., the robustness of the results from the statistical analyses based on the FAS.

17.1. PRIMARY EFFICACY

17.1.1. PRIMARY EFFICACY VARIABLE & DERIVATION

The primary efficacy endpoint is progression-free survival (PFS), which is defined as the time from the date of randomization to the earliest of documented disease progression (as defined by RECIST v 1.1 per investigator

radiology review) or death from any cause. Disease assessment will continue until disease progression, even after the originally assigned treatment is discontinued. Subjects for whom progression or death do not occur will be censored according to rules described in Section 17.1.2.

The primary efficacy variable is comprised of PFS times which are calculated after applying the censoring rules that were specified for each analysis. PFS time will be calculated (in months) as (the earliest date of progression, death, or censoring, minus the randomization date, plus 1) divided by 365.25 days per year, times 12 months per year. PFS in tables and listings will be displayed to the closest 1/10 of a month.

17.1.2. Missing Data Methods for Primary Efficacy Variable

The analyses of PFS in this study are designed to be consistent with FDA guidance documents and the EMA scientific guidelines for analysis of PFS.

The EMA Guideline on the Evaluation of Anticancer Medicinal Products in Man (CHMP/EWP/205/95 Rev 34), Appendix 1, suggests that the primary analysis of PFS (defined as death or documented disease progression) should follow the ITT principles and be based on actual death dates and actual date of documentation of disease progression.

PFS Censoring Rules (Set 1) are designed to match this guidance except that subjects who receive non-protocol anticancer treatment without death or documentation of disease progression beforehand will be censored at the date of last disease assessment prior to the start of non-protocol anticancer treatment per May 2007 FDA Guidance for Industry "Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics".

The guidelines also suggest that sensitivity analyses should be conducted to address important assumptions in the methods used, including the handling of deviations and missing data, censoring for subsequent anti-cancer treatment, non-radiological disease progression, and the handling of unscheduled evaluations. The PFS Censoring Rules (Sets 2, 3, 4) described in the sections below are designed to produce sensitivity analyses which address some of these concerns.

Under each set of censoring rules, if a subject fulfils the criteria for more than one situation, then the most conservative applicable outcome ("progressed", followed by "censored") will be assigned to the subject. Also, the shortest applicable time to event within the assigned outcome category will be used for the analysis.

17.1.2.1. PFS Censoring Rules (Set 1)

The PFS Censoring Rules (Set 1) will be used to manage missing data in the primary analysis of PFS. Under the Set 1 rules, PFS will be right-censored for subjects who meet one or more of the following conditions:

- Subjects with no post-baseline disease assessments (unless death occurred prior to the first scheduled assessment, in which case death will be considered a PFS event)
- Subjects who initiate subsequent anticancer therapy or intervention in the absence of documented progression
- Subjects who die or have disease progression after missing 2 or more consecutively scheduled disease assessment visits
- Subjects who are last known to be alive and progression-free on or before the data cut-off date

For such subjects, PFS will be right-censored according to the conventions described in Table E. These conventions are based on the May 2007 FDA Guidance for Industry "Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics".

Table E. PFS Censoring Rules (Set 1): Date of Progression or Censoring

Situation	Date of Progression or Censoring	Outcome
Death before first scheduled disease assessment	Actual date of death	Progressed
Documented disease progression before first scheduled disease assessment	Actual date of disease assessment at which progression was noted	Progressed
Death or documented disease progression between 2 scheduled disease assessments	Actual date of death or actual date of first disease assessment showing documented disease progression, whichever occurs first	Progressed
Missed ≥2 consecutive disease assessments immediately prior to death or progression	Actual date of last non-missing disease assessment that is before the missed visits and prior to death and prior to documented disease progression	Censored
No death and no post-baseline disease assessments	Actual date of randomization	Censored
Non-protocol anticancer treatment started prior to death or prior to documentation of disease progression	Actual date of last non-missing disease assessment prior to start of non-protocol anticancer treatment	Censored
Alive and not progressed [1]	Actual date of last non-missing disease assessment	Censored

[1] Without death and without documentation of disease progression; includes subjects who may have discontinued from study prior to death and without documentation of disease progression.

17.1.2.2. PFS Censoring Rules (Set 2)

PFS Censoring Rules (Set 2) will be used for sensitivity analyses to address potential bias related to the violation of the assumption of non-informative censoring and missing disease assessments. The definition of a PFS event is expanded from the conventional definition of PFS specified in Set 1 rules to include subjects who discontinue protocol treatment, start non-protocol treatment prior to death or progression, or miss ≥2 consecutive disease assessments immediately prior to death or progression.

Under the Set 2 rules, the occurrence of a PFS event will be imputed (i.e. not be considered censored) under the following conditions:

- Discontinuation of treatment will be considered a failure of therapy. Therefore, subjects who discontinue therapy without death and without documented disease progression will be counted as an observed PFS event at the time of discontinuation.
- Patients for whom documentation of disease progression or death occurs immediately after ≥2 consecutive missed disease assessments will be considered to have had an observed PFS event on the date of the last disease assessment visit without documentation of disease progression prior to the missed visits.
- Patients who use any non-protocol anticancer treatment prior to documented disease progression will be considered to have had an observed PFS event on the start date of the new therapy.

PFS will be right-censored for subjects who meet one or more of the following conditions:

- Subjects with no post-baseline disease assessments unless death occurred prior to the first scheduled assessment (in which case death will be considered a PFS event).
- Subjects who are last known to be alive and progression-free on or before the data cut-off date.

Table F summarizes the handling of missing data and PFS determinations under the Censoring Rules (Set 2).

Table F. PFS Censoring Rules (Set 2): Date of Progression or Censoring

Situation	Date of Progression or Censoring	Outcome
Death before first scheduled disease assessment	Actual date of death	Progressed
Death or documented disease progression between 2 scheduled disease assessments	Actual date of death or actual date of first disease assessment showing documented disease progression, whichever occurs first	Progressed
Missed ≥2 consecutive disease assessments immediately prior to death or progression	Actual date of last non-missing disease assessment visit without a documented PFS event that is before the missed visits and before death or progression	Progressed
Treatment discontinuation prior to death and prior to documented disease progression	Actual date of the decision to discontinue treatment	Progressed
Non-protocol anticancer treatment started prior to death or prior to documented disease progression	Actual date of new treatment start	Progressed
No death and no post-baseline disease assessments	Actual date of randomization	Censored
Alive and no PFS* event	Actual date of last disease assessment	Censored

**The definition of PFS event under Set 2 rules includes the situations categorized as "Progressed" in the table.*

17.1.2.3. PFS Censoring Rules (Set 3)

The PFS Censoring Rules (Set 3) are based on Table B (PFS 2 – uniform progressions and assessment dates) of the May 2007 FDA Guidance for Industry "Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics". Set 3 rules are designed to correct for potential bias due to any differences in the follow-up schedules for disease assessments, including unscheduled visits. These alternative censoring rules will be used for sensitivity analyses of the PFS.

Under Set 3 rules, an observed PFS event is defined as documented disease progression or death due to any cause. The occurrence of a PFS event will be imputed to the date of the scheduled disease assessment time. Accordingly, the dates of observed events will be adjusted as follows:

- If disease progression is documented between 2 scheduled disease assessments, the planned date of the later scheduled assessment will be assigned as the date of progression.
- In the event of death, the date of the endpoint will not be adjusted.

- If the response at a scheduled time (X) is classified as Indeterminate (IND) or Not Assessed (NA) and at the next scheduled timepoint (X+1) the best response is PD, then the planned date of the first timepoint (X) will be used as the uncensored date of progression.
- Censoring will be similar to Set 1, except that the planned dates of scheduled assessments will be used for the censored values.

Table G summarizes the handling of missing data and PFS determination under the Censoring Rules (Set 3).

Table G. PFS Censoring Rules (Set 3): Uniform Progression and Assessment Dates

Situation	Date of Progression or Censoring	Outcome
Death before first scheduled disease assessment	Actual date of death	Progressed
Death or documented disease progression at a scheduled disease assessment	Actual date of death or planned date of the scheduled disease assessment, whichever occurs first	Progressed
Death or documented disease progression between scheduled disease assessments	Actual date of death or planned date of next subsequent scheduled disease assessment after the documented disease progression, whichever occurs first	Progressed
A disease assessment of "indeterminate" or "not assessed" at Visit X, followed by documented disease progression at Visit X+1	Planned date of scheduled disease assessment at Visit X	Progressed
Non-protocol anticancer treatment started prior to death or without prior documentation of disease progression	Planned date of last non-missing scheduled disease assessment prior to start of non-protocol anticancer treatment.	Censored
No death and no post-baseline disease assessments	Actual date of randomization	Censored
Missed ≥2 consecutive disease assessments immediately prior to death or progression	Planned date of the first of >=2 missed disease assessment visits prior to death and prior to documented disease progression	Censored
Alive and no PFS event*	Planned date of last non-missing scheduled disease assessment	Censored

**The definition of PFS event under Set 3 rules includes the situations categorized as "Progressed" in the table.*

17.1.2.4. PFS Censoring Rules (Set 4)

The PFS Censoring Rules (Set 4) are designed to address possible limitations of the scientific methods used for documentation of disease progression by including symptomatic or non-radiological disease progression as an observed PFS event.

Set 4 rules will be used for sensitivity analyses. Under Set 4 rules, an observed PFS event is defined as clinical (i.e., symptomatic or non-radiological) disease progression, documented disease progression, or death due to any cause. Subjects who do not experience clinical or documented disease progression or death during the analysis period will be censored.

PFS will be right-censored for subjects who meet one or more of the following conditions:

- Subjects with no post-baseline disease assessments (unless death occurred prior to the first scheduled assessment, in which case death will be considered a PFS event)
- Subjects who initiate subsequent anticancer therapy or intervention in the absence of documented progression
- Subjects who die or have disease progression immediately progression immediately after missing 2 or more consecutively scheduled disease assessment visits
- Subjects who are last known to be alive and progression-free on or before the data cut-off date

For such subjects, PFS will be right-censored according to the conventions described in Table H.H These conventions are based on the May 2007 FDA Guidance for Industry "Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics".

Table H. PFS Censoring Rules (Set 4): Date of Progression or Censoring

Situation	Date of Progression or Censoring	Outcome
Death before first scheduled disease assessment	Actual date of death	Progressed
Death, documented disease progression or clinical progression between 2 scheduled disease assessments	The earliest of: (1) actual date of death, (2) actual date of first disease assessment showing clinical disease progression, (3) actual date of first disease assessment showing documented disease progression	Progressed
Missed ≥2 consecutive disease assessments immediately prior to a PFS event	Actual date of last non-missing disease assessment that is prior to the missed visits and prior to any PFS event	Censored
No death and no post-baseline disease assessments	Actual date of randomization	Censored
Non-protocol anticancer treatment started prior to any PFS event	Actual date of last non-missing disease assessment prior to start of non-protocol anticancer treatment	Censored
Alive and no PFS* event	Actual date of last non-missing disease assessment	Censored

**The definition of PFS event under Set 4 rules includes the situations categorized as "Progressed" in the table.*

17.1.3. PRIMARY ANALYSIS OF PRIMARY EFFICACY VARIABLE

The primary objective of this study is to test the hypothesis that PFS is superior (longer) for subjects treated with AGS-16C3F than for subjects treated with axitinib.

The primary analysis of PFS based on the investigator's assessment of disease will be performed in accordance with the ITT principle in the FAS. All subjects who undergo randomization will be included in the primary analysis and grouped according to the treatment groups assigned by randomization. The Censoring Rules (Set 1) which are presented in Section 17.1.2.1 will be used to derive the event status and PFS times.

The difference between treatment groups will be tested using a stratified log rank test at a one-sided 0.1 significance level. The stratification factors will be the ECOG PS and the number of prior systemic RCC regimens, without any other covariate(s). If the direction of the difference between treatment arms favors AGS-16C3F, then the observed significance level (p-value) of the 1-sided test of superiority of AGS-16C3F will be calculated as 0.5 * p-value from a two-sided log-rank test (corresponding to a chi-squared test with d.f.=1). If the direction of the difference favors axitinib, then the observed significance level will be calculated as (1 – (0.5 * p-value from the two-sided log-rank test)).

Kaplan-Meier survival curves for the primary analysis variable will be produced, overall and by stratum. The following descriptive statistics will be reported in a table, overall and by stratum:

- Progression status:
    - Number and percent of subjects with observed disease progression per RECIST
    - Number and percent of subjects who died without disease progression beforehand
    - Number and percent of censored subjects
- Number and percent of censored subjects, by reason for censoring
- Estimated progression-free survival time (months), including the minimum, 25th percentile, 50th percentile (median), 75th percentile, and maximum. The 2-sided 95% CIs for the 25th, 50th, and 75th percentiles, calculated using Greenwood's formula for the standard error, will also be reported.

The overall adjusted hazard ratio (HR) for AGS-16C3F relative to axitinib will be estimated using a proportional hazards (PH) model that includes terms for treatment effect and a stratification factor based on the baseline ECOG PS and the number of prior systemic RCC regimens (see section 7.1.1), and no other covariate. Ties will be addressed using Efron's (1977) method.

The model-based estimates for the HR and its corresponding one-sided 90% CI (upper confidence bound) will be reported. A two-sided 95% CI will also be reported in the clinical study report.

Hazard ratio estimates by stratum will be obtained from a PH model containing treatment, the stratification factor, and a treatment-by-stratum- interaction term. Homogeneity in the hazard ratios among strata will be examined by Wald's test. Results from a PH model without stratification will be reported also.

The adequacy of the model(s) will be evaluated, including an assessment of the proportional hazards assumption, as described in Section 17.1.6, "Exploratory Analyses".

17.1.4. SECONDARY ANALYSES OF PRIMARY EFFICACY VARIABLE

Sensitivity analyses will be performed to evaluate the robustness of the PFS results derived under the primary analysis (Bhattacharya 2009, Carroll 2007, and Stone 2011). Such sensitivity analyses will assess potential sources of bias due to 1) differences between the planned and actual schedule of disease assessments, 2) censoring for subsequent anticancer therapy, 3) symptomatic or non-radiological disease progression, and 4) missed disease assessments due to reasons such as protocol deviations, consent withdrawal, and subjects lost to follow-up.

17.1.4.1. Sensitivity to Analysis Set

Methods used for the primary analysis of the primary PFS endpoint will be used in the PPS to investigate the sensitivity to analysis set (i.e., sensitivity to major protocol deviations) and robustness of the primary analysis results. The Censoring Rules (Set 1) apply.

17.1.4.2. Sensitivity to Missing Data Assumptions

The Censoring Rules (Set 2) are designed to address the assumption of non-informative censoring and are presented in Section 17.1.2.2. After application of the Censoring Rules (Set 2), the same methods that were used in the primary analysis will be applied in the FAS population.

The Censoring Rules (Set 3) are designed to correct for potential bias due to any differences in the follow-up schedules for disease assessments (Section 17.1.2.3). After application of the Censoring Rules (Set 3), the same methods that were used in the primary analysis will be applied in the FAS population.

17.1.4.3. Sensitivity to Limitations in Scientific Methods

The Censoring Rules (Set 4) are designed to address possible limitations of the scientific methods used for documentation of disease progression by including symptomatic or non-radiological disease progression as an observed PFS event. After application of the Censoring Rules (Set 4), the same methods that were used in the primary analysis will be applied in the FAS population.

17.1.5. Subgroup Analyses of Primary Efficacy Variable

This study is not powered for statistical testing by subgroup. The interpretive focus of subgroup analyses will be on the descriptive statistics, point estimates, and confidence intervals by subgroup, rather than any within-group statistical tests.

The consistency of treatment effect across subgroups of subjects will be examined by repeating the primary analyses with replacement of the randomization stratification factor by a single planned covariate and its interaction with the treatment group variable. The planned covariates include age (<=65, >65), gender (male, female), race (Caucasian, not Caucasian), ECOG Performance Status at baseline (0, 1), histology (clear cell, non-clear cell), number of prior treatment regimens for RCC (2 for clear cell or 1 for non-clear cell, >2 for clear cell or >1 for non-clear cell), and Heng's prognostic risk group (favorable, intermediate, poor). The subgroup levels of each covariate were defined in Section 7.6.

Hazard ratio estimates by subgroup will be obtained from a Cox PH model containing treatment, the subject characteristic of interest (i.e., the "covariate"), and a treatment-by-covariate interaction term. The model-based hazard ratio estimates for each level of the covariate and the adjusted hazard ratio (stratified by ECOG and number of prior regimens for RCC) from the primary analysis (main effect model), along with their respective two-sided 95% CI's, will be displayed on a Forest plot.

Subgroup analyses of secondary endpoints and/or secondary determinations are not planned.

17.1.6. Exploratory Analyses of Primary Efficacy Variable

Section 17.4, "Subgroup Analyses", describes exploratory analyses related to the possible effects of selected covariates.

Analyses and plots which are generated for the purpose of validating statistical assumptions are also considered exploratory, and may be included or excluded from the clinical study report appendices at sponsor discretion.

Planned exploratory investigations of this nature include investigations of the censoring patterns and adequacy of PH models for the study data.

The assumption of non-informative censoring will be examined through sensitivity analyses with censoring rules sets 2 and 4. The censoring pattern will be assessed by reviewing the distribution of censoring times by treatment arm for differences between groups. Also, the cumulative percentage of censored subjects will be plotted against the observed censoring times, by treatment arm.

The models which are proposed for examination in this study include a (1) a PH model of treatment, stratified jointly for 2 randomization factors (primary analysis); (2) an unstratified PH model containing only a term for treatment; (3) a PH model containing terms for treatment and one covariate; and (4) PH models containing terms for treatment, a covariate, and treatment-by-covariate interaction. A total of 7 covariates will be examined; these are described in Section 17.4. Each covariate will be considered a time-fixed covariate (i.e., the values of the covariate cannot be written as a function of the survival time), categorical variable. If a PH model is appropriate, the hazard function h(t|x) at time t given a vector of time-fixed covariate(s) will be proportional to the baseline hazard function $h_0(t)$ at t and the model will have a general form $h(t|x) = h_0(t)\, e^{(\alpha + x\beta)}$, where t is a subject's survival time, x is the vector of covariate(s) associated with the subject, h_0 is the baseline hazard function for an individual with x=0 for continuous variables and reference cell values of any categorical variable .

The proportional hazard assumption of the planned models will be assessed using suggestions from Therneau and Grambsch (2000) and Lawless (1982), as follows.

- The scaled Schoenfeld residuals for a covariate over event time will be calculated under each model. A plot of the scaled Schoenfeld residuals against event time should have a zero slope if the PH assumption reasonably holds for this covariate.
- A plot of log negative log survival function of the lifetimes versus the log of the lifetimes (i.e., log [-log $S_0(t|x_i)$] versus log t) within each level of x_i should produce roughly parallel lines if a Cox PH model is appropriate.

Further exploratory investigations of departures from the planned models may be conducted if suggested by the data. Exploratory analyses that are defined after database lock will be described in a separate requirements document.

17.2. SECONDARY EFFICACY

Each of the secondary endpoints described below will be analyzed in the FAS and PPS:

- Progression-free survival (PFS), based on blinded central radiology assessment
- Overall Survival (OS)
- Objective Response (OR), based on the investigator's radiographic assessment
- Disease control (DC), based on the investigator's radiographic assessment
- Duration of Objective Response (DOR), based on the investigator's radiographic assessment

Three additional endpoints will be analyzed in the FAS and PPS to investigate sensitivity to the disease assessment methodology:

- OR based on blinded central radiology assessment
- DC based on blinded central radiology assessment
- DOR based on blinded central radiology assessment

17.2.1. SECONDARY EFFICACY VARIABLES & DERIVATIONS

17.2.1.1. Progression-Free Survival Based on Blinded Central Radiology Assessment

Progression-free survival based on the blinded central radiology assessment is defined as the time from the date of randomization to the earliest of documented disease progression as defined by RECIST v 1.1, as recorded by the central radiologist, or death from any cause. Subjects for whom death or progression are not observed will be censored.

PFS will be calculated in months as (the earliest date of progression, death, or censoring, minus the randomization date, plus 1) divided by 365.25 days per year, times 12 months per year. PFS in tables and listings will be displayed to the closest 1/10 of a month.

17.2.1.2. Overall Survival

Overall survival is defined as the time from the date of randomization until the date of death from any cause. Subject death information is collected on the Serious Adverse Event and Survival Follow-Up eCRFs.

Subjects for whom death is not observed will be censored using the Censoring Rules (Set 5) presented in Table I (Section 17.2.2.2). OS will be calculated in months as (the earliest date of death or censoring, minus the randomization date, plus 1) divided by 365.25 days per year, times 12 months per year. OS in tables and listings will be displayed to the closest 1/10 of a month.

17.2.1.3. Objective Response Based on Investigator's Radiographic Assessment

The primary determination of objective response (OR) is based on the investigator's radiographic assessment of tumor response. OR is defined as a best overall response of Complete Response (CR) or Partial Response (PR). A best response of CR or PR must be confirmed no less than 28 days after the criteria for response are first met.

17.2.1.4. Disease Control Based on Investigator's Radiographic Assessment

The primary determination of disease control (DC) will be based on the investigator's radiographic assessment of tumor response. DC is defined as a best overall response of Complete Response (CR), Partial Response (PR), or at least 6 months of Stable Disease (SD) after the date of randomization.

17.2.1.5. Duration of Objective Response Based on Investigator's Radiographic Assessment

The primary determination of duration of objective response (DOR) will be based on the investigator's radiographic assessment of tumor response. The DOR is defined as the time in months from the date of the first response of CR or PR, whichever is first recorded and subsequently confirmed, to the earlier of the date of first documented progressive disease or the date of death due to any cause. Subjects for whom death or progression are not observed will be censored.

DOR will be calculated using the date of first documented progressive disease, death, or censoring as the end date of the objective response. The Censoring Rules (Set 1) from Section 17.1.2.1 will be applied. The calculation of DOR will be the date of the end of objective response minus the date of first confirmed response plus 1 day, divided by 365.25 days per year, times 12 months per year. DOR in tables and listings will be displayed to the closest 1/10 of a month.

17.2.1.6. OR, DC, and DOR Based on Blinded Central Radiology Assessment

The OR, DC, and DOR will be derived using the blinded central radiography assessment of tumor response. Definitions and derivations are otherwise the same as described for the OR, DC, and DOR based on the investigator's assessment of tumor response (Section 17.2.1.3, Section 17.2.1.4, and Section 17.2.1.5, respectively).

17.2.2. Missing Data Methods for Secondary Efficacy Variables

17.2.2.1. Progression-Free Survival Based on Blinded Central Radiology Assessment

The Censoring Rules (Set 1) which were presented in Section 17.1.2.1 will apply to PFS based on the blinded central radiology assessment of tumor response.

17.2.2.2. Overall Survival

Subjects who have not died will be right-censored as shown in Table I.

Table I. Overall Survival Censoring Rules (Set 5): Date of Progression or Censoring

Situation	Date of Death or Censoring	Outcome
Any death	Actual date of death	Death
No death and no post-baseline contact	Actual date of randomization	Censored
No death and at least 1 post-baseline-up	Actual date of last subject contact [1]	Censored

[1] The latest of all post-baseline assessment dates where the subject was physically alive, e.g. disease assessment visit, blood draw, non-fatal AE start/end date, dosing visit, date last reported alive, etc.

17.2.2.3. Objective Response Based on Investigator's Radiographic Assessment

If the best overall response is missing or not evaluable, the subject will be classified as not having an objective response and will only be counted in the denominator for calculating ORR.

17.2.2.4. Disease Control Based on Investigator's Radiographic Assessment

If the best overall response is missing, the subject will be classified as not achieving disease control and will only be counted in the denominator for calculating DCR.

17.2.2.5. Duration of Objective Response Based on Investigator's Radiographic Assessment

Subjects who experience a response and do not experience death or progression after that response will have a censored PFS time. The Censoring Rules (Set 1) will be applied prior to calculating the duration of response.

17.2.2.6. OR, DC, and DOR Based on Blinded Central Radiology Assessment

Imputation of missing values for the OR, DC, and DOR based on the blinded central radiology review will be the same as for the OR, DC, and DOR based on the investigator assessment. The censoring rules (Set 1) will be applied to DOR based on central radiology review.

17.2.3. ANALYSES OF SECONDARY EFFICACY VARIABLES

All secondary efficacy variables will be analyzed in the FAS and PPS.

17.2.3.1. Progression-Free Survival by Blinded Central Radiology Assessment

The analysis methods and models that were used for the primary analysis will be applied to the PFS assessment by the blinded central radiology reviewer, using Censoring rules Set 1.

17.2.3.2. Overall Survival

Overall survival will be analyzed using the same statistical methodology as described for PFS based on the Blinded Central Radiology assessment in Section 17.2.3.1, except that Censoring Rules Set 5 will apply and the statistical tests will be 2-sided, conducted at a 0.05 level of significance.

17.2.3.3. Objective Response Rate Based on Investigator's Radiographic Assessment

A Cochran-Mantel-Haenszel analysis of the ORR, stratified for baseline ECOG-PS and number of prior systemic therapies for RCC, will be conducted at a two-sided significance level of 0.05. The stratified Mantel-Haenszel estimate of the odds ratio for experiencing objective response, with the axitinib arm as the reference level, will be reported as a measure of relative treatment effect, along with its two-sided 95% CI. Homogeneity among the 4 strata will be examined using a two-sided Breslow-Day test, conducted at a significance level of 0.05.

The ORR will also be summarized with estimates for each treatment group, overall and by stratum, based on the crude proportion of subjects who have a best overall response of CR or PR. The Clopper-Pearson method will be used to calculate exact two-sided 95% CIs for each crude rate.

17.2.3.4. Disease Control Rate Based on Investigator's Radiographic Assessment

The disease control rate (DCR) will be analyzed using the same methods as for ORR (Section 17.2.3.3).

17.2.3.5. Duration of Objective Response Based on Investigator's Radiographic Assessment

The duration of response as assessed by the investigator will be analyzed using the same statistical methodology as described for PFS based on the Blinded Central Radiology assessment in Section 17.2.3.1. . Subjects who do not experience an objective response and therefore have no duration of objective response will be excluded from the analyses.

17.2.3.6. OR, DC, and DOR Based on Blinded Central Radiology Assessment

The OR, DC, and DOR based on blinded central radiology reviewer's assessment will be analyzed using the same statistical methods as the OR, DC, and DOR based on the investigator's assessment, respectively.

17.2.3.7. Concordance of Disease Assessments by Independent Reviewers

The best overall response, objective response, and disease control determinations made by the investigator will be compared to the consensus determinations of the blinded central reviewers. Additionally, determinations made by the two radiologists who conducted the blinded central review will be compared.

These concordance analyses will be performed by treatment group and overall. For each analysis, the number and percentage of subjects for whom the assessors were in agreement will be presented.

17.3. OTHER EFFICACY VARIABLES

17.3.1. OTHER EFFICACY VARIABLES & DERIVATIONS

Tumor response and its components (target lesion response, non-target lesion response, timepoint response, and best overall response) were not clearly stated as analysis endpoints in the protocol; however, these measurements support the planned analysis endpoints and the protocol specified some supportive summaries and analyses which are described in this section.

Briefly, a disease assessment based on CT scan and/or MRI is planned every 8 weeks (±7 days), counting from C1D1. Tumor response is determined using RECIST v 1.1 and recorded in the Disease Assessment eCRF. Best overall response is collected once per subject, on the Subject Status - Completion eCRF. A best response of CR or PR must be confirmed no less than 28 days after the criteria for response are first met.

Tumor response and best overall response are determined independently by the investigator and the central radiology reviewers. The central radiology review consists of assessment by 2 independent reviewers, followed by adjudication of the results if the 2 reviewers disagree. The primary and all secondary efficacy endpoints are based on these tumor response assessments plus any recorded subject death.

[Note: Clinical symptoms of disease progression are recorded as adverse events. Bone imaging scans and brain scans provide additional documentation of disease.]

17.3.2. MISSING DATA METHODS FOR OTHER EFFICACY VARIABLES

No imputation of missing data will be performed for tumor response or best overall response.

17.3.3. ANALYSES OF OTHER EFFICACY VARIABLES

Tumor response (investigator and consensus of the blinded central radiology reviewers) will be summarized by treatment group and visit. The number of subjects who were assessed at each visit will be reported. The number and percentage of subjects who had each possible response will be tabulated for the target lesions, the non-target lesions, and the (timepoint) overall response.

Best overall response (investigator and consensus of the blinded central radiology reviewers) will be tabulated for each treatment group using the crude proportion of subjects whose best response during the course of protocol treatment is Complete Response (CR), Partial Response (PR), Stable Disease (SD), Progressive Disease (PD) and non-evaluable (NE). Subjects with a rating of NE will be included in the denominator.

Swimmer lane plots will be produced for FAS subjects, showing duration of treatment, start of complete response, start of partial response, response episode end, continuing response (if applicable), and durable responders.

A waterfall plot showing the maximum reduction of sum of diameters of target lesions and the best overall response will be produced for the FAS.

18. QUALITY OF LIFE ANALYSIS

Quality of life is assessed using the ECOG Performance Status (ECOG PS) at Screening, at the end of each Cycle (e.g., approximately every 3 weeks) for the AGS-16C3F subjects, and at each planned visit for the axitinib subjects.

The ECOG-PS will be summarized by treatment group with a shift table from baseline to each subject's worst post-baseline value and from baseline to final assessment in the study. The analyses will be performed in the FAS and PPS.

19. SAFETY OUTCOMES

All analyses and outputs for safety outcomes will be based on the Safety Analysis Set (SAF). There will be no statistical comparisons between the treatment groups for safety data, unless otherwise specified within the relevant section.

19.1. ADVERSE EVENTS

Treatment emergent adverse events (TEAEs) are defined as AEs that started or worsened in severity after the first dose of study drug and until 28 days post the last administration of the study drug or the subject's Safety Follow-Up visit, whichever occurred later. Events that begin on Day 1 will be counted as treatment-emergent.

AEs will be coded to system organ class (SOC) and preferred term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA) Version 18.1 central coding dictionary and will be graded according to the National Cancer Institute (NCI) Common Toxicity Criteria for Adverse Events (CTCAE), version 4.03.

All summaries of AEs will include only treatment-emergent events unless otherwise stated.

Safety will be evaluated based on the incidence of AEs (including Grade 3 and Grade 4 AEs, treatment-related adverse events, serious adverse events, and AEs requiring discontinuation of study treatment). The number and percentage of AEs, SAEs, AEs leading to discontinuation of study treatment, and AEs related to study drug will be summarized by SOC, PT and treatment group. The number and percentage of AEs by severity will also be summarized. Listings of AEs, SAEs, deaths, and events leading to discontinuation of study treatment will be presented.

See Appendix 2 for handling of partial dates for AEs. In the case where it is not possible to define an AE as treatment emergent or not, the AE will be classified by the worst case; i.e. treatment emergent.

An overall summary of the number and percentage of subjects who had events in each of the categories described below will be produced by treatment group and for combined treatment groups.

- Subjects with at least one AE
- Subjects with at least one possibly drug-related AE
- Subjects with at least one AE of CTCAE Grade 3 or Grade 4
- Subjects with at least one possibly drug-related AE of CTCAE Grade 3 or Grade 4
- Subjects with at least one serious AE
- Subjects with at least one possibly drug-related serious AE
- Subjects whose treatment had to be modified due to an AE
- Subjects whose treatment had to be modified due to a possibly drug-related AE
- Subjects who discontinued treatment with study drug because of an AE
- Subjects who discontinued treatment with study drug because of a possibly drug-related AE
- Subjects with at least one AE resulting in death

- Subjects with at least one possibly drug-related AE resulting in death

Listings will include TEAEs and Non-TEAEs. Any AE that was reported with a start date on or after the initiation of non-protocol anticancer treatment will be flagged in the listings.

19.1.1. ALL TEAEs

Incidence of TEAEs will be presented by treatment group, System Organ Class (SOC), and Preferred Term (PT), and will also be broken down further by maximum severity and relationship to study drug. In these tables, SOCs will be presented alphabetically, and PTs will be presented by decreasing frequency of all events in a PT in the AGS-16C3F arm within SOC.

The number and percentage of subjects who reported TEAEs, possibly treatment-related TEAEs, and TEAEs of CTCAE Grade ≥3 will also be summarized by treatment group and presented by PT in the order of descending incidence in the AGS-16C3F group.

19.1.1.1. Severity

Severity will be assigned by the investigator according to the CTCAE severity grades. The severity of each event will be recorded as Grade 1 (mild), Grade 2 (moderate), Grade 3 (severe), Grade 4 (life-threatening) or Grade 5 (fatal).

For any interim tables, TEAEs starting after the first dose of study drug with a missing severity will be left missing in the summary tables and identified in the data issues log(s). If the severity of a TEAE is still missing at the time of database lock, the event severity will be left missing and footnoted on the severity table. If a subject reports a TEAE more than once within that SOC/PT, the AE with the worst case severity will be used in the corresponding severity summaries.

19.1.1.2. Relationship to Study Drug

Relationship, as indicated by the Investigator, is collected in response to the question, “Is there a reasonable possibility that the event may have been caused by the study drug? (Yes, No)”. A response of “Yes” to this question indicates a related TEAE; a response of “No” indicates an unrelated TEAE.

TEAEs with a missing relationship to study drug will be tabulated as missing. If a subject reports the same AE more than once within that SOC/PT, the AE with the strongest relationship to study drug will be used in the corresponding relationship summaries.

19.1.2. TEAEs LEADING TO DISCONTINUATION OF STUDY DRUG

TEAEs leading to permanent discontinuation of study drug will be identified by using the “Action Taken” field on the Adverse Event eCRF, based on selection of the text value “study drug discontinued”.

A summary of the incidence of TEAEs leading to discontinuation of study drug by SOC and PT will be prepared. A listing of TEAEs that led to discontinuation of study drug will be provided.

19.1.3. SERIOUS ADVERSE EVENTS

Serious adverse events (SAEs) are those events recorded as “Serious” on the Adverse Events page of the eCRF. A summary of serious TEAEs by SOC and PT will be prepared. A listing of SAEs will be provided.

19.1.4. GRADE 3/GRADE 4 TEAEs

A summary of Grade 3 and Grade 4 TEAEs by treatment group, SOC, PT, and toxicity grade will be provided. Grade 3 and Grade 4 events will also be listed. [Note: In some places, labels will say "Grade 3 or Grade 4", depending on the particular usage.]

19.1.5. ADVERSE EVENTS LEADING TO DEATH

TEAEs leading to Death are those events which are recorded on the Serious Adverse Events page of the eCRF as having an outcome of death. A summary of TEAEs leading to death by SOC and PT will be prepared. A listing of AEs that led to death will be provided, based on data from the SAE eCRF.

19.1.6. ADVERSE EVENTS OF SPECIAL INTEREST

Ocular adverse events and diagnoses which are entered as AEs will be reported on the general tables of AEs. A listing of eye-related TEAEs will be provided.

Ocular toxicities based on ophthalmology exams will be reported on the Eye Exam eCRFs and will be graded using the study specific ocular criteria. These will be tabulated separately from the general AE tables. See Section 19.7.1 and Section 19.7.2 for planned analysis and reporting.

19.1.7. CTCAE GRADING OF ADVERSE EVENTS

AEs will be graded using the Common Terminology Criteria for Adverse Events as defined in the CTCAE 4.03 (14 June 2010) grading system. The complete reference document may be found here:

http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf

Appropriate details will be included in the programming specifications documents.

19.2. DEATHS

Adverse events that led to death are addressed in Section 19.1.5.

This study also collects mortality information on the Survival Follow-Up form, which does not collect a cause of death. All deaths due to any cause during the study will be part of the overall survival analyses (Section 17.2).

19.3. LABORATORY EVALUATIONS

Local laboratories will be used in this study. Results from blood chemistry, hematology, coagulation, urine tests (including parameters calculated from a 24-hour urine collection), and thyroid tests are expected.

No lab data imputation will be performed. Incomplete data may be excluded from some analyses. In cases where an analysis depends on identifying a clinically significant abnormal value, only values which are non-missing and which can be clearly identified programmatically as significantly abnormal will be counted.

19.3.1. LABORATORY PARAMETERS

A list of laboratory assessments to be included in the outputs is included below.

19.3.1.1. Hematology

- Red blood cells
- Hematocrit
- Hemoglobin
- Platelets
- White blood cells
- Absolute neutrophils
- Absolute eosinophils
- Absolute basophils
- Absolute lymphocytes
- Absolute monocytes

19.3.1.2. Serum Chemistry

- Sodium
- Potassium
- Chloride
- Bicarbonate
- Glucose
- BUN
- Creatinine
- Total protein
- Albumin
- Total Bilirubin
- Direct bilirubin
- ALT (SGPT)
- AST (SGOT)
- Alkaline phosphatase
- LDH
- Uric acid
- Calcium
- Magnesium

- Phosphorus (Phosphate)
- C-reactive protein (CRP)

19.3.1.3. Urine Tests

- Urine protein to creatinine ratio (uPCR)

19.3.1.4. 24-Hour Urine Collection

- Urine collection time
- Urine collection volume
- Urine rate of production in ml/hour (derived)
- 24-hour protein

19.3.1.5. Coagulation

- Prothrombin (PT)
- Activated partial thromboplastin time (aPTT)
- International normalized ratio (INR)

19.3.1.6. Thyroid Function Tests

- TSH
- Free T4

19.3.2. Laboratory Specific Derivations

- Unit conversions for laboratory data will be provided in the programming specifications documents.

19.3.3. Laboratory Reference Ranges and Markedly Abnormal Criteria

Quantitative laboratory measurements will be compared with the relevant laboratory reference ranges in SI units and categorized as:

- Low: Below the lower limit of the laboratory reference range
- Normal: Within the laboratory reference range (upper and lower limit included)
- High: Above the upper limit of the laboratory reference range
- Laboratory values that do not have a reference range will not be categorized.

19.3.4. CTCAE Grading for Laboratory Data

Laboratory values for 18 analytes (23 CTCAE terms) will be graded by computer, using the Common Terminology Criteria for Adverse Events (CTCAE) grading system v4.03 (14 June 2010). Table J provides details.

Table J. CTCAE Criteria for Evaluation of Laboratory Values

EDC Collection	CTCAE Term	
Analyte	**High**	**Low**
Albumin		Hypoalbuminemia
Alkaline Phosphatase	Alkaline phosphatase increased	
ALT (SGPT)	Alanine aminotransferase increased	
AST (SGOT)	Aspartate aminotransferase increased	
Bilirubin, Total	Blood bilirubin increased	
Calcium	Hypercalcemia	Hypocalcemia
Creatinine	Creatinine increased	
Glucose	Hyperglycemia	Hypoglycemia
Magnesium	Hypermagnesemia	Hypomagnesemia
Phosphorus		Hypophosphatemia
Potassium	Hyperkalemia	Hypokalemia
Sodium	Hypernatremia	Hyponatremia
aPTT	Activated partial thromboplastin time prolonged	
WBC		White blood cell decreased
HGB	Hemoglobin increased	Anemia
PLT		Platelet count decreased
Absolute Neutrophils		Neutrophil count decreased
Absolute Lymphocytes	Lymphocyte count increased	Lymphocyte count decreased

The complete reference document may be found here:

http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf

It is expected that incomplete data will result in a missing CTCAE Grade.

19.3.5. Laboratory Analysis Methods

Presentations will use SI Units. Quantitative laboratory measurements reported as "<X", i.e. below the lower limit of quantification (BLQ), or ">X", i.e. above the upper limit of quantification (ULQ), will be converted to X for the purpose of quantitative summaries, but will be presented as recorded, i.e. as "<X" or ">X" in data listings.

The following summaries by treatment arm will be provided for laboratory data:

- Incidence of Grade 3 and Grade 4 toxicities, based on the CTCAE grading system, at each subject's worst post-baseline value, safety follow-up assessment, and final assessment
- Shift tables from baseline toxicity grade to worst post-baseline toxicity grade, based on the CTCAE grading system

Listings of all laboratory data collected during the study will be produced. Any positive pregnancy tests will be listed separately.

19.4. ECG EVALUATIONS

19.4.1. ECG ASSESSMENTS

Subjects will have an ECG at screening, on C1D1, and at the end of study safety follow-up visit. Subjects who are randomized to AGS-16C3F will have an ECG at predose and at 2 hours post-dose on C1D1; subjects randomized to receive axitinib will have an ECG at predose only, on C1D1.

ECGs are done in triplicate for all timepoints, approximately 3-5 minutes between each assessment. Each set of results will be identified by a unique recorded clock time. The triplicate values may be considered individually or represented by their median, depending on the analysis. Retests and unscheduled assessments will be included, unless otherwise specified.

The following ECG parameters will be reported for this study:

- HR (bpm)
- PR Interval (msec)
- QRS Interval (msec)
- QT Interval (msec)
- QTcF Interval (msec) [derived]
- Overall assessment of ECG (Investigator's judgment):
  - Normal
  - Abnormal, Not Clinically Significant (ANCS)
  - Abnormal, Clinically Significant (ACS)

Any ECG-related adverse events that are determined by the investigator or medical reviewer to have occurred will be collected on the Adverse Event CRF and reported in the AE tables, listings, and figures.

19.4.2. ECG SPECIFIC DERIVATIONS

- Fridericia's Correction (msec)
  - $QTcF\ (\text{msec}) = \frac{QT\ (ms)}{\sqrt[3]{RR\ (ms)/1000}}$

- RR Interval – If RR Interval is not available for the derivation of the QTc corrections, it will be derived from HR as

- $RR\ (\text{msec}) = 1000 * \frac{60}{HR\ (bpm)}$

All components for these calculations must be taken from the same assessment, which will be identified in the database by the date-time of the ECG recording.

19.4.3. CTCAE Grading of QT Interval

The Common Terminology Criteria for Adverse Events (CTCAE) grading system v4.03 (14 June 2010) will be used to grade possible ECG QT corrected interval prolongation. The complete reference document may be found here:

http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf

The initial computerized grading will be based on the QTcF numeric values only, excluding any clinical component. The initial computerized grades will be assigned as shown in Table K and reported on the ECG results listing:

Table K. Criteria for Evaluation of ECG QT corrected interval prolongation

CTCAE Term	CTCAE Grade	One or more of the following conditions:
QTcF prolongation	1	The median of the triplicate measurements at a planned timepoint is within (450 ms ≤ QTcF ≤480 ms)
	2	The median of the triplicate measurements at a planned timepoint is within (481 ms ≤ QTcF ≤500 ms)
	3	Any 2 of the triplicate measurements are in (501 ≤ QTcF)
	4	The median of the triplicate measurements is within (501 ≤ QTcF) or the median change from baseline among the triplicate measurements is (QTcF change from baseline >60 ms).

[Note: If only 1 measurement is available, the single measurement will be taken as the "median".]

Grade 3 and Grade 4 values do not automatically qualify as adverse events; there is also a clinical component to determining whether an AE has occurred. For example, a subject who has a median QTcF >= 501ms or a median change from baseline >60 ms will be flagged as having a (possible) Grade 4 event according to the computerized criteria. However, the final determination of a Grade 4 event depends on clinical considerations (e.g., occurrence of Toursade de pointes or polymorphic ventricular tachycardia or signs/symptoms of serious arrhythmia).

Therefore, the investigator and/or other medical reviewer will consider the subject's complete data before making a final determination about recording an ECG-related AE and assigning the final CTCAE grade. The investigator has the final responsibility to record and grade any AE that may have occurred.

19.4.4. ECG Analysis Methods

The ECG quantitative data will be summarized by incidence of treatment-emergent Grade 3 or Grade 4 QTcF values by treatment arm. Grade 3 and Grade 4 values will be identified prior to averaging the triplicate ECG results.

Subjects who met the abnormal criterion specified in Section 19.4.3 will be identified on the ECG data listing.

19.5. VITAL SIGNS

Vital signs will be recorded once at each visit for subjects who are receiving axitinib. They will be recorded at predose and at 30 minutes post-dose for subjects who are receiving AGS-16C3F. The following vital signs will be collected.

- Sitting Systolic Blood Pressure (mmHg)
- Sitting Diastolic Blood Pressure (mmHg)
- Sitting Pulse Rate (bpm)
- Respiratory Rate (breaths/min)
- Temperature (^{o}C)

Predose body weight (kg) will also be recorded once at each clinic visit for all subjects.

19.5.1. VITAL SIGN SPECIFIC DERIVATIONS

If required, subject body temperature will be converted from Fahrenheit to Centigrade using the formula,

- °C = 5/9 x (°F - 32)

If required, subject body weight will be converted from pounds (lbs) to kilograms (kg) using the formula,

- kg = lbs/2.2

19.5.2. CTCAE GRADING OF CHANGES IN BODY WEIGHT

Percent changes in subject body weight will be evaluated against the CTCAE criteria presented in Table L.

Table L. Criteria for Evaluation of Changes in Subject Body Weight

CTCAE Term	CTCAE Grade	One or more of the following conditions:
Weight Gain	1	+5% to < +10% change from baseline
	2	+10% to < +20% change from baseline
	3	≥ +20% change from baseline
Weight Loss	1	-5% to > -10% change from baseline
	2	-10% to > -20% change from baseline
	3	≤ -20% change from baseline

19.5.3. VITAL SIGN ANALYSIS METHODS

The maximum weight increase from baseline and maximum weight decrease from baseline (per subject), expressed as an absolute change and a percent change, will be summarized using descriptive statistics. The number and percentage of subjects who met CTCAE criteria for weight gain (Grade 1, 2, or 3) compared to baseline, or weight loss (Grade 1, 2, or 3) compared to baseline, will be presented.

Vital signs will be listed; values which met the CTCAE criteria for Grade 1, 2, or 3 weight gain or weight loss will be identified on the listing.

Clinically important treatment-emergent vital sign values or changes from baseline may lead to the identification and recording of an adverse event.

19.6. PHYSICAL EXAMINATIONS

Post-baseline abnormalities will be reported as adverse events. Physical examination results will be listed by treatment group.

19.7. OTHER SAFETY ASSESSMENTS

19.7.1. OPHTHALMOLOGIC EXAMS

A complete eye exam is planned to be performed at Baseline for all subjects, at end of Cycle 2, end of Cycle 6, and at Safety Follow-up for AGS-16C3F subjects, and at other times if clinically indicated for any subject. Results will be collected on the Eye Exam eCRF page. The following types of data will be collected.

- Ophthalmology Symptoms
- Visual Acuity
- Extraocular Movement
- Tonometry
- External Exam
- Slit Lamp Exam
- Dilated Fundus Exam
- Optional Exams
- Impression
- Visual Acuity Grade
- Ocular Diagnoses and Ophthalmology Outcome, including the diagnosis, event status (new, ongoing, resolved), relatedness to study drug, prescribed treatment, and action taken with study treatment due to the eye exam results.

In general, eye-related summaries by timepoint will include data from both planned and PRN assessments. Determinations at each timepoint will include all eye data collected since the previous planned assessment. The screening period will include eye assessments performed within 28 days prior to the first dose of study drug. For subjects who receive AGS-16C3F, Day 43 will include the worst value recorded for each subject from post-dose

C1D1 through Predose C3D1, inclusive. Day 106 will include the worst value recorded from Post-dose C3D1 through C6D1, inclusive. Subsequent intervals will be defined similarly at 3-cycle intervals during treatment. For subjects who receive axitinib, Day 43 will include the worst value recorded for each subject from Day 1 through Day 43 (±7 days). Day 106 will include the worst value recorded from the Day 43 assessment date + 1 through Day 106 (±7 days), inclusive. Subsequent intervals will be defined similarly at 63-day intervals during treatment. Data collected after the decision to end treatment for AGS-16C3F subjects or after the date of last dose for axitinib subjects are considered to be "end of treatment and beyond".

Ocular symptoms recorded during the complete eye exam will be coded using the MedDRA Version 18.1 dictionary. The first post-baseline occurrence of any ocular symptom and of each ocular symptom will be summarized over time. The number of subjects at risk in each time interval, the adjusted conditional probability of experiencing the symptom for the first time in an interval, the cumulative probability of experiencing the specified symptom for the first time from the time of first dose through the end of the interval, and the total number of reports of the symptom in the specified interval will be presented in a table. The number of reports will include multiple reports per subject where applicable.

Detailed visual acuity measurements (e.g., 20/20 or 20/40, etc.) and Investigator's Impression will be listed but not summarized.

Visual acuity changes from baseline, represented by the number of lines of change from baseline (see Table M), will be summarized by treatment arm and assigned ocular AE grade, by eye and overall for both eyes. The investigator's determination of changes in visual acuity and the ocular AE grade of any changes will be used in all analyses. The number and percentage of subjects who had a change which qualified as a Grade 1, 2, 3, or 4 change will be summarized at each planned timepoint, for the worst post-baseline value, and at the final assessment. [Note: Subjects are not expected to improve from baseline. Cases of improvement will be returned to Data Management for investigation.]

Abnormalities of extraocular movement, pupil exam, tonometry (IOP), external exam, slit lamp exam (cornea, lens, lids and lacrimation, conjunctive and sclera, anterior chamber, iris, vitreous, other), and dilated fundus exam (disc, disk/cup ratio, macula, vessels, periphery) will be summarized at each planned visit, at the worst value, and at the final assessment, by the number and percentage of subjects having abnormalities in each eye (where applicable) and in both eyes. In the case of intraocular pressure, IOP >21 mmHg in an eye will be considered abnormal (Chang DF 2011). If both the right eye and the left eye meet this criterion, then "both eyes" will be considered abnormal. The categories of "abnormal at baseline", "abnormal, same as previous", and "abnormal, new or change from previous" will be derived for the IOP. In the case of the dilated fundus exam, statistics will be generated for exams performed with dilation and without dilation.

Results from optional exams (corneal pachymetry, Schirmer, corneal topography, endothelial evaluation, and other diagnostics) will be listed.

Ocular diagnoses will be coded using the MedDRA Version 18.1 dictionary. The number and percentage of subjects who had any ocular diagnosis during the study and the number and percentage who had each unique diagnosis (based on the PT) will be reported by treatment arm. The statistics will be provided for all diagnoses and also for drug-related diagnoses.

Ocular diagnoses will be reported (along with the investigator's Visual Acuity Grade and the data from the Ophthalmologic Outcome" section of the eCRF) as ocular adverse events. The reporting of these events is addressed in Adverse Events (Section 19.1). Grading of these events is presented in Section 19.7.1.1.

Complete eye exam data as collected will be presented by subject in a data listing. No data will be imputed.

19.7.1.1. Derivations Related to Grading of Ocular AEs

The grading of Ocular AEs is based on changes in visual acuity (VA delta). In this study, the VA may be measured with or without correction, and it should always be non-missing. The same method which is used to measure VA at baseline (uncorrected or corrected) should be used throughout the study on all eye exams for an individual subject.

The 2 measurement methods (corrected, uncorrected) must not be mixed in making the determination of the VA delta. In cases where multiple methods are used within a subject in the study, the baseline measurement method will determine which of the post-baseline assessments can be used to assign a VA delta, and which should result in a VA delta of "cannot assess". If no VA delta can be assigned, then the VA Grade will be missing. For example, if a subject's data includes a baseline exam VA (uncorrected), a post-baseline exam 1 VA (uncorrected), and post-baseline exam 2 VA (corrected), then a VA delta and VA Grade can be assigned to only the post-baseline exam 1 result. The post-baseline exam 2 result can not be assessed for VA delta, and the associated VA Grade would be missing.

In cases where VA was not done or measured consistently for a subject but BCVA was done consistently throughout, the BCVA data may be used for VA grading. For example, if VA is missing at baseline, the baseline BCVA and post-baseline BCVA measurements may be used to assess changes in visual acuity. However, this is a compromise to the original study plan, as BCVA is a more lenient measure of change. The methods of assigning the VA delta and VA Grade are otherwise the same as when VA is used (as described above).

Non-standard visual acuity data should be converted to standard values prior to analysis. Visual acuity measurements which include special characters such as '+' or '-', with or without an associated number, should be stripped of the special characters and the numeric part of the value should be used as is. In addition, the numeric part of the value should be converted to a standard value, if necessary, by rounding down. Table M provides some examples.

Details of the grading process follow.

Table M. Conversion of Visual Acuity Data from Non-Standard to Standard Values

Source data	Derived Value for Analyses
20/32	20/30
20/30 +	20/30
20/30 +2	20/30
20/30 -	20/30
20/30 -2	20/30
20/30 -2 +2	20/30

Protocol Table 12 describes the grading of ocular AEs based on changes from baseline in visual acuity (VA) or best corrected visual acuity (BCVA). It is reproduced for convenience in Table N. Cases in which a subject's baseline vision was recorded as better than 20/20 will be grouped with the 20/20 line of Table N for the purpose of assessing post-baseline changes.

Table N. Visual Acuity Grading Chart

Baseline vision [1]	Grade 1 [2]	Grade 2 [3]	Grade 3 [3]	Grade 4
20/20	20/20	20/25 (1) 20/30 (2) 20/40 (3)	20/50 (4) 20/60 (5) 20/70 (6) 20/80 (7) 20/100 (8) 20/125 (9)	20/200 or worse
20/25	20/25	20/30 (1) 20/40 (2) 20/50 (3)	20/60 (4) 20/70 (5) 20/80 (6) 20/100 (7) 20/125 (8)	20/200 or worse
20/30	20/30	20/40 (1) 20/50 (2) 20/60 (3)	20/70 (4) 20/80 (5) 20/100 (6) 20/125 (7)	20/200 or worse
20/40	20/40	20/50 (1) 20/60 (2) 20/70 (3)	20/80 (4) 20/100 (5) 20/125 (6)	20/200 or worse
20/50	20/50	20/60 (1) 20/70 (2) 20/80 (3)	20/100 (4) 20/125 (5)	20/200 or worse
20/60	20/60	20/70 (1) 20/80 (2) 20/100 (3)	20/125 (4)	20/200 or worse
20/70	20/70	20/80 (1) 20/100 (2) 20/125 (3)	N/A	20/200 or worse
20/80	20/80	20/100 (1) 20/125 (2)	N/A	20/200 or worse
20/100	20/100	20/125 (1)	N/A	20/200 or worse
20/125	20/125	N/A	N/A	20/200 or worse
20/200	20/200	N/A	N/A	Worse than 20/200

[1] With or without correction, based on patient's daily functional distance vision
[2] No change in vision from baseline. Grade 1 is the same as screening (baseline) from a VA perspective. The Grade 1 Clinical Ocular AE will be an event when the AE is present (e.g., keratopathy) but there is no change in VA from baseline and the subject has little to no symptoms. If the AE is not treated aside from using artificial tears, then it would be considered Grade 1.

[3] Visual acuity measurement and delta, where delta = number of line changes from baseline.

Clinical Ocular AE grading will be determined based on visual acuity (VA) change from screening (baseline). This will be done as follows:

- Determine baseline VA for each eye. Subject should have the VA measured with or without glasses depending on the subject's requirement for daily use for distance vision. Use Table L and identify the row of the baseline vision for each eye.
- Subsequent eye exam VA (obtained under same conditions as screening (baseline)), determine if a change in vision is present. If so, use the row in Table L identified in Step 1, to locate the current VA and determine the number of line-of-vision-change from baseline for each eye; this is expressed as the number in parenthesis. Also note the grade of the VA for each eye.
- Between the two eyes, determine the eye with greater line-of-vision-change from baseline (delta). The Clinical Ocular AE grade for the subject is the grade of the worse (delta) eye. See Examples 1 and 2 in Table M.
- In the event both eyes have the same delta, then the Clinical Ocular AE grade comes from the eye with the worse new vision. See Example 3 in Table O.

Table O. Examples of Clinical Ocular AE: Keratopathy

	Baseline (CC)	New VA (CC)	Lines VA Change	VA Grade	Clinical Ocular AE Grade
Example 1					
OD	20/20	20/50	(4)	3	Grade 3
OS	20/40	20/70	(3)	2	
Example 2					
OD	20/60	20/125	(4)	3	Grade 3
OS	20/100	20/200	(2)	4	
Example 3					
OD	20/30	20/100	(5)	3	Grade 4
OS	20/60	20/200	(5)	4	

19.7.2. EYE SYMPTOM QUESTIONNAIRE

The Eye Symptom Questionnaire (ESQ) will be completed at screening and on other occasions when the full ophthalmologic exam is performed. The ESQ collects information related to the following items:

- Dry Eyes (none, mild, moderate, severe)
- Eye Pain (none, mild, moderate, severe)

- Blurred Vision (none, mild, moderate, severe)
- Photophobia (none, mild, moderate, severe)
- Other, specify (none, mild, moderate, severe)
- Steroid Use (Taken as directed, used most of the time, did not use most of the time, used every day continuously, did not use at all, other)

A summary table by treatment arm will present the number and percentage of subjects who reported each eye symptom, overall and by severity, during each cycle and at the Safety Follow-up visit. Each unique symptom report will be counted only in the cycle in which it began. The severity of the worst post-baseline assessment and the final assessment will also be tabulated. A data listing will be provided.

A separate table will display the number and percentage of subjects who had >25%, >50%, >75%, or 100% of ocular symptom-free days during treatment. The calculated number and percentage of symptom-free days by subject will be provided in a listing.

Also, a swimmer lane plot which shows the duration of any post-baseline ocular symptom during the period from Day 1 through the end of the Safety Follow-up period will be produced. The plot will display a bar connecting any consecutive visits at which a symptom was reported, or a simple bracket or narrow line around individual visits at which a symptom was reported. Important treatment events (e.g., day of first dose, day of last dose or decision to end treatment, day of Safety Follow-up visit, day of each study drug infusion for AGS-16C3F subjects) will be marked on the subject's timeline.

A summary of ocular steroid use by cycle and at the final assessment will be produced.

20. PHARMACOKINETIC ANALYSIS

The Sponsor pharmacokineticist will perform the analyses described in Protocol Section 7.6 and will provide the plasma concentration data and PK parameter estimates to QuintilesIMS.

Descriptive statistics for plasma concentration data will be provided by visit and planned timepoint. Descriptive statistics for the PK parameter estimates will be provided by visit.

Plasma concentration data and PK parameter estimates will be listed.

21. ANTI-DRUG ANTIBODY

Blood samples for testing anti-AGS-16C3F antibody level will be collected from subjects who receive AGS-16C3F on C1D1, C4D1, C6D1, and every 12 weeks thereafter until treatment is discontinued.

The number and percent of subjects (incidence) of anti-drug antibody formation to human native antibody (AGS-16C3) and to antibody drug conjugate (AGS-16C3F) will be tabulated at each timepoint and at each subject's final assessment.

22. DATA NOT SUMMARIZED OR PRESENTED

These domains and/or variables will not be summarized or presented on listings, but will be available in the clinical study database, SDTM and/or ADaM datasets, or in the non-computerized study report appendices.

Subject data

- Comments
- Demographics: If female, is the subject a woman of childbearing potential? (Y/N)
- AGS-16C3F drug administration: Infusion start time, infusion stop time, infusion rate, infusion unit
- Axitinib drug administration data: Dispensing and return information, miscellaneous other.
- Survival status: If alive or unknown, will this subject's survival follow-up continue? (Y/N). If no, please indicate reason (Lost to follow-up/ Withdrew consent). Date lost to follow-up or date consent withdrawn
- SAEs: Date of death, cause of death, date of hospital admission/discharge, study drug administration lot number, start/stop/ongoing, dose, unit, frequency (as collected on the SAE eCRF).
- Details of specialized informed consent, including:
    - Pre-Screen Informed Consent
    - Subject consent to optional tumor sample collection as part of SOC while on study
    - Subject consent to bank leftover tumor tissue samples for future research testing to learn more about the study drug
    - Subject consent to bank leftover tumor tissue samples for future research testing to explore how to treat cancer
- Randomization number

<u>CRF fields designed primarily for data management purposes or as CRF completion instructions</u>

- CRF fields designed to track what procedures were performed or what CRFs should have been completed.
    - Indicate what unscheduled CRFs were completed
    - Indicate if chemistry tests were performed? (Y/N)
    - Indicate if hematology tests were performed? (Y/N)
    - Indicate if coagulation tests were performed? (Y/N)
    - Indicate if urine tests were performed? (Y/N)
    - Was a serum pregnancy test done? (Y/N)
    - Indicate if PK blood draw was performed? (Y/N)
    - Indicate if ADA blood draw was performed? (Y/N)
    - Was bone imaging performed? (Y/N)
    - Was a brain scan performed? (Y/N)
- CRF fields designed to account for the presence/absence of other recorded data on a CRF.
    - Did subject meet all protocol eligibility criteria? (Y/N)
    - Was any surgery performed for malignancy prior to enrollment? (Y/N)
    - Was any radiotherapy given for malignancy prior to enrollment? (Y/N)
    - Physical examination (normal/abnormal) and related instructions for recording further data on other CRFs
    - Ocular diagnoses at this visit? (Y/N)

- Events CRF

- CRF fields designed to facilitate the tracking of collected samples or follow-up activities.
  - PK Requisition number
  - ADA requisition number
  - SAE: Date Agensys notified of report
  - SAE: Site awareness date
  - SAE: Is a hospital discharge summary available? (Y/N)
  - SAE: Was an autopsy performed? (Y/N)
- CRF fields designed to help track subject participation.
  - Will the subject continue another dosing cycle? (Y/N)

23. REFERENCES

American Heart Association, "Understanding Blood Pressure Readings". Retrieved 20 Oct 2015. http://www.heart.org/HEARTORG/Conditions/HighBloodPressure/AboutHighBloodPressure/Understanding-Blood-Pressure-Readings_UCM_301764_Article.jsp#.VibIU36rTrd

AJCC Cancer Staging Manual, 7th edition (2010). Editors: Edge S, Byrd D, Compton CC Greene FL, Trotti A. Springer-Verlag New York.

Bhattacharya S, Fyfe G, Gray RJ, Sargent DJ (2009). Role of sensitivity analyses in assessing progression-free survival in late-stage oncology trials. Journal of Clinical Oncology 27:5958–5964.

Carroll KJ (2007). Analysis of progression-free survival in oncology trials: Some common statistical issues. Pharmaceutical Statistics 6:99–113.

Chang DF (2011). Ophthalmologic examination. In P Riordan-Eva, ET Cunningham, eds., Vaughan and Asbury's General Ophthalmology, 18th ed., pp. 27-57. New York: McGraw-Hill.

Dobson, A. J. (2002). *An introduction to generalized linear models, 2nd ed.* Chapman & Hall/CRC.

Efron, B. (1977), "The Efficiency of Cox's Likelihood Function for Censored Data," *Journal of the American Statistical Association*, 72, 557–565.

Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J (2009). New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). European Journal of Cancer 45: 228-247.

Ellenberg SS, Eisenberger MA(1985). An efficient design for phase III studies of combination chemotherapies. Cancer Treatment Reports 69: 1147–1154.

EMA. (24 June 2010.). Guideline on Missing Data in Confirmatory Clinical Trials. Retrieved December 15, 2011, from European Medicines Agency. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/09/WC500096793.pdf

EMA (13 December 2012). Guideline on the Evaluation of Anticancer Medicinal Products in Man (CHMP/EWP/205/95 Rev 4) Appendix 1. "Methodological consideration for using progression-free survival (PFS) or disease-free survival (DFS) in confirmatory trials", from European Medicines Agency. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/01/WC500137126.pdf

Heng DYC, Wanling X, Regan MM, Warren MA, Golshayan AR, Sahi C, Eigl BJ, Ruether JD, Cheng T, North S, Venner P, Knox JJ, Chi KN, Kollmannsberger C, McDermott DF, Oh WK, Atkins MB, Bukowski RM, Rini BI and Choueiri TK (2009). "Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: Results from a large, multicenter study". American Society of Clinical Oncology.

Lawless JF (1982) *Statistical Models and Methods for Lifetime Data*, pages 365-366. John Wiley and Sons.

Oken, M.M., Creech, R.H., Tormey, D.C., Horton, J., Davis, T.E., McFadden, E.T., Carbone, P.P. (1982). Toxicity And Response Criteria Of The Eastern Cooperative Oncology Group. Am J Clin Oncol 5:649-655. The Eastern Cooperative Oncology Group, Robert Comis M.D., Group Chair.

Schwartz LH, Litie`re S, de Vries E, Ford R, Gwyther S, Mandrekar S, Shankar L, Bogaerts J, Chen A, Dancey J, Hayes W, Hodi FS, Hoekstra OS, Huang EP, Lin N, Liu Y, Therasse P, Wolchok JD, Seymour L. RECIST 1.1 Update and clarification: From the RECIST committee. Euro J Cancer (2016) 62:132-137.

Stone AM, Bushnell W, Denne J, Sargent DJ, Amit O, Chen C, Bailey-Iacona R, Helterbrand J, Williams G (2011). Research outcomes and recommendations for the assessment of progression in cancer trials from a PhRMA

working group. European Journal of Cancer 47:1763–1771.

Therneau TM, Grambsch PM (2000). Modeling Survival Data: Extending the Cox Model. Springer New York.

U.S. Code of Federal Regulations, 21CFR101.74 (revised as of April 1, 2015). Retrieved 20 Oct 2015.

U.S. Department of Health and Human Services National Institutes of Health, National Cancer Institute: Common Terminology Criteria for Adverse Events (CTCAE) v4.03 (June 14, 2010).

U.S. Food & Drug Administration: Guidance for Industry. Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. May 2007.

U.S. NIH Library of Medicine [MedLinePlus Encyclopedia] "Vital Signs", 01/21/2013. Retrieved 20 Oct 2015. https://www.nlm.nih.gov/medlineplus/ency/article/002341.htm

Company Documents

Protocol AGS-16C3F-15-3: A Multi-center, Open Label, Randomized Phase 2 Study of AGS-16C3F vs. Axitinib in Metastatic Renal Cell Carcinoma, Version 4.0 (Amendment 3), dated 18 Jul 2017. IND 113,865. Agensys, Inc.

AGS-16C3F-15-3 Subject Case Report Forms, Version 6.0 (dated 26 SEP 2017).

AGS-16C3F-15-3 Data Management Interim Database Lock Plan

AGS-16C3F-15-3 Unblinding Plan (dated 21 SEP 2017)

APPENDIX 1. QUINTILES GLOBAL BIOSTATISTICS STANDARD OUTPUT CONVENTIONS

ABBREVIATIONS

ASCII	American standard code for information interchange file format
CGM	Computer graphics metafile
ODS	Output Delivery System
RTF	Rich text file format

INTRODUCTION

This Appendix applies to standards used for outputting tables, listings and figures. It is intended to provide specifications to guide the statistician or statistical programmer in setting up specifications for programming tables, listings and figures. These standards should be used in the absence of customer specific standards.

If the mock shells for this study appear to depart significantly and consistently from these standards in any way, please check with the project statistician for a resolution prior to programming.

OUTPUT FILE NAMING CONVENTIONS

- File names should only consist of uppercase letters, lowercase letters, digits (0 to 9) and underscores. A period should only be used to indicate a separator between the file name and the extension. No spaces, other special characters or punctuation marks are permitted.
- As far as possible, output files should be in RTF format, although .DOC files are also permitted. The program, program log and output file name should reflect the type and number of the statistical output. If this is not possible, then the output name should be at least as descriptive as possible.
- A prefix can be used to distinguish between a Table, Listing and Figure document ('T' for table, 'L' for listing and 'F' for figure). If there is only 1 digit in the number of the table, listing or figure in the place where 2 digits are possible, a leading zero should be added in the file name to make sorting consistent with the sequence (e.g., T14_3_01_1.RTF).

PAPER SIZE, ORIENTATION AND MARGINS

- The size of paper will be Letter for the United States, otherwise A4.
- The page orientation should preferably be landscape, but portrait is also permitted.
- Margins should provide at least 1 inch (2.54 centimeters) of white space all around the page, regardless of the paper size.
- The number of columns per page (linesize) should be 145 for A4 and 134 for Letter.
- The number of rows per page (pagesize) should be 49 for A4 and 51 for Letter.

FONTS

- The font type 'Courier New' should be used as a default for tables and listings, with a font size of 8. The font color should be black. No bolding, underlining italics or subscripting should be permitted. Try to avoid using super-scripts, unless absolutely necessary. Single spacing should be used for all text.
- Figures should have a default font of "Times Roman", "Helvetica", or "Courier New".
- This can be achieved by using the following options in SAS.
    - goptions
    - gunit = pct
    - cback = white
    - colors = (black)
    - hby = 2.4
    - ftext = "TimesRoman"
    - htext = 2.5
    - device = cgmof97l
    - gaccess = gsasfile;
    - filename gsasfile "....cgm";

HEADER INFORMATION

Headers should be defined as follows:

- The header should be placed at the top of the page (same place on each page) regardless of the size or orientation of the table or listing.
- The customer name should appear in row 1, left-aligned.
- The output page number (page x of y) should appear in Row 1, right-aligned.
- The protocol number should appear in Row 2, left-aligned
- The output identification number should appear in row 3, centered.
- The output title should start in row 4, centered.
- The output population should appear in row 5, centered. The population should be spelled out in full, e.g., Intention-to-Treat in preference to ITT and should be placed in parentheses.
- Row 6 should be a blank line.
- Mixed case should be used for titles.
- The output titles should be designed so that they are arranged consistently through all outputs. For example, content (e.g., Vital Signs) followed by metric (e.g., Vital Signs Change from Baseline).
- Titles should not contain quotation marks or footnote references.
- The column headings should be underlined with a row of underscores ('_').
- Column headings spanning more than one column should be underlined and have underscores on either side of the title and should be centered.
- Column headings containing numbers should be centered.
- Column headings should be in sentence case.

- In general, the population count should appear in the column header in the form "(N=XXX)".
- "Statistic" should be the column header over n, Mean, SE, n (%) etc.
- As a rule, all columns should have column headings.

TABLE AND LISTING OUTPUT CONVENTIONS

General

- The first row in the body of the table or listing should be blank.
- The left hand column should start in column 1. No indenting or centering of the output should occur.
- Rounding should be done for display purposes, after all calculations have been completed, using the SAS function ROUND.
- Numbers in tables should be rounded, not truncated.
- Alphanumeric output should be left aligned.
- Numbers should be decimal point aligned.
- Whole numbers should be right aligned.
- Text values should be left aligned.
- The first letter of a text entry should be capitalized.
- Listings of adverse events, concomitant medications, medical histories etc. should be sorted in chronological order, with earliest adverse event, medication or history coming first.
- The study drug should appear first in tables with treatments as columns.
- In general, only present totals (across treatment groups) at baseline/randomization, and do not present them post randomization, unless the customer specifically requests it.
- If possible, include 100% frequencies in the table shell, so that it is clear what the denominator is for percentage calculations.
- All listing outputs should be sorted (preferably by Treatment, Site Number and Subject Number).
- Do not use superscripts and subscripts.
- Exponentiation will be expressed using a double asterisk, i.e., mm3 will be written as mm**3.
- All variables that are output in the CRF (which have data present) should appear in the listings, along with all derived data appearing in the corresponding tables.
- The width of the entire output should match the line size.

Univariate Statistics

- Statistics should be presented in the same order across tables (i.e., n, Mean, SD, Median, Minimum, Maximum).
- Table statistics should line up under the N part of the (N=XXX) in the table header. All decimal points should line up. If the minimum and maximum are output on one line as Minimum, Maximum then the comma should line up with the decimal point.

- If the original data has N decimal places, then the summary statistics should have the following decimal places.
    - Minimum and maximum: N
    - Mean, median and CV%: N + 1
    - SD: N + 2

Frequencies and percentages (n and %)

- Percent values should be reported inside parentheses, with one space between the count and the left parenthesis of the percentage. Parentheses should be justified to accept a maximum of 100.0 as a value and padded with blank space if the percent is less than 100.0. An example is given below.
    - 77 (100.0%)
    - 50 (64.9%)
    - 0 (0.0%)
- Percentages will be reported to one decimal place. Percentages that are <100.0% but >99.9% will be presented as '>99.9%' (e.g., 99.99% is presented as >99.9%). Percentages that are <0.1% will be presented as '<0.1%' (e.g., 0.08% is presented as <0.1%). Rounding will be applied after the <0.1% and >99.9% rule.
- Percentages may be reported to 0 decimal places as appropriate (for example, where the denominator is relatively small).
- Where counts are zero, percentages of 0.0% should appear in the output.

Confidence Intervals

- As a rule confidence intervals are output to one place more than the raw data, and standard deviations and standard errors to two places more than the raw data.
- Confidence intervals should be justified so that parentheses displayed on consecutive lines of a table "line up".
- Boundary values of confidence intervals should be separated by a comma.
- Boundary values should be padded as necessary to accept negative values and to allow alignment of the decimal place.
- An example is given below:
    - (-0.12, -0.10)
    - (9.54, 12.91)

P-values

P-values should be reported to three decimal places, except values <1.000 but >0.999 will be presented as '>0.999' (e.g., 0.9998 is presented as >0.999); and values <0.001 will be presented as '<0.001' (e.g., 0.0009 is presented as <0.001). Rounding will be applied after the <0.001 and >0.999 rule.

Ratios

- Ratios should be reported to one more decimal place than the original data.

Spacing

- There must be a minimum of 1 blank space between columns (preferably 2).

Denominators

- If a different count other than the population count is used for a denominator (within the table) to calculate percentages, there should be a row in the table that identifies that number "n".

- Alternatively, a footnote should be included in each table with percentages to indicate the denominator for percentages.

Missing values

- A "0" should be used to indicate a zero frequency.
- A blank will be used to indicate missing data in an end-of-text table or subject listing.

FIGURE OUTPUT CONVENTIONS

- Figures should be provided in RTF files using the SAS Output Delivery System (ODS), as Computer Graphics Metafile (CGM) formatted graphical output generated by SAS.
- The CGM file itself should contain the title or footer.
- The image should be clear and of high quality when viewed in the Word document, and when printed.
- In general, boxes around the figures should be used.

FOOTNOTE INFORMATION

Footers should be defined as follows:

- A continuous line of underscores ('_') will follow the body of the table or listing prior to any footnotes at the bottom of the page.
- Table footnotes should be defined using compute statements in the proc report, and should appear directly after the body of the table.
- The program path and name and version number (if applicable) should appear as footnote 1 at the bottom of the page.
- The date/time stamp should appear as footnote 2 at the bottom of the page.
- Footnotes should be left-aligned.
- Footnotes should be in sentence case.
- Only "typewriter" symbols are permitted – e.g., "*", "$", "#", "@", "&" and "+".
- The choice of footnote symbols should be consistent. E.g. if you have the footnote "# indicates last observation carried forward" for one table, the same symbol and footnote should indicate LOCF for all tables.
- If text wraps across more than one line (for a note), the first letter for all lines of text after the first one will be indented to align beneath the first letter of the text in the first line.
- The page identification in the format Page X of Y (where Y is the total number of pages for the output) should appear in the first footnote, right aligned.
- Ordering of footnotes should be as follows:
    - 1.) Source data listing reference, if necessary
    - 2.) Abbreviations and definitions
    - 3.) Formulae
    - 4.) P-value significance footnote

- 5.) Symbols
- 6.) Specific notes

- Common notes from table to table should appear in the same order.
- The symbols should appear in the same order as what they are defined in the table or listing, from left to right.

PROGRAMMING INSTRUCTIONS

Programming instructions must appear in blue font at the end of each table or listing shell. Programming instructions, where necessary, should follow the table or listing shells in blue font, beginning with the words "Programming Note" followed by a colon. These include notes on the output, reminders of how to handle missing values, repeat shells for similar tables etc.

DATES & TIMES

Depending on data available, dates and times will take the form "yyyy-mm-dd hh:mm".

SPELLING FORMAT

English US

PRESENTATION OF TREATMENT GROUPS

For outputs, treatment groups will be represented as follows and in the order shown.

Treatment Group	For Tables and Graphs	For Listings (include if different to tables)
AGS-16C3F	AGS-16C3F	
Axitinib	Axitinib	
Total (where applicable)	Total	Not applicable

PRESENTATION OF TIMEPOINTS

For outputs, treatment cycles will be represented as follows and in the order shown. Visits will be displayed as carried in the raw data, unless otherwise specified in the analysis dataset specifications. The number of visits and cycles will vary across subjects.

Long Name	Short Name (default for presentations)
Cycle x Day y	CxDy
Example:	
Cycle 1 Day 1	C1D1
Cycle 1 Day 2	C1D2

LISTINGS

All listings will be ordered by the following (unless otherwise indicated in the template):

- Randomized treatment group (or treatment received if it's a safety output)
- Subject ID
- Date (where applicable)

APPENDIX 2. PARTIAL DATE CONVENTIONS

ALGORITHM FOR TREATMENT EMERGENCE OF ADVERSE EVENTS

Imputed dates will not be presented in the listings.

Start Date	Stop Date	Action
Known	Known	If start date < study med start date, then not TEAE If study med start date ≤ AE start date ≤ the later of (study med stop +28 days or the safety follow-up visit), then TEAE If the later of (study med stop +28 days or the safety follow-up visit) < AE start date, then not TEAE
	Partial	If start date < study med start date, then not TEAE If study med start date ≤ AE start date ≤ the later of (study med stop +28 days or the safety follow-up visit), then TEAE If the later of (study med stop +28 days or the safety follow-up visit) < AE start date, then not TEAE
	Missing	If start date < study med start date, then not TEAE If study med start date ≤ AE start date ≤ the later of (study med stop +28 days or the safety follow-up visit), then TEAE If the later of (study med stop +28 days or the safety follow-up visit) < AE start date, then not TEAE
Partial, but known components show that it cannot be on or after study med start date	Known	Not TEAE
	Partial	Not TEAE
	Missing	Not TEAE
Partial, could be on or after study med start date	Known	If stop date < study med start date, then not TEAE If study med start ≤ AE stop date ≤ the later of (study med stop +28 days or safety follow-up visit), then TEAE If the later of (study med stop +28 days or safety follow-up visit) < AE stop date, then TEAE

Start Date	Stop Date	Action
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study med start date, then not TEAE If study med start ≤ AE stop date ≤ the later of (study med stop +28 days or the safety follow-up visit), then TEAE If the later of (study med stop +28 days or the safety follow-up visit) < AE stop date, then TEAE
	Missing	Assumed TEAE
Missing	Known	If stop date < study med start date, then not TEAE If study med start ≤ AE stop date ≤ later of (study med stop +28 days or the safety follow-up visit date), then TEAE If the later of (study med stop +28 days or safety follow-up visit date) < AE stop date, then TEAE
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study med start date, then not TEAE If study med start ≤ AE stop date ≤ the later of (study med stop +28 days or safety follow-up visit date), then TEAE If the later of (study med stop +28 days or safety follow-up visit date) < AE stop date, then TEAE
	Missing	Assumed TEAE

ALGORITHM FOR PRIOR/CONCOMITANT/POST MEDICATIONS

Note: Separate flags will be assigned for prior, concomitant, and "post" medications. It is possible for a medication to be considered a prior medication, a concomitant medication, or both, depending on when the medication was used. However, only medications that began post-treatment will be flagged as "post".

Note: Although a prior medication flag is defined here, prior medications will not be summarized in tables.

Rule	Rule Type	Start Date	Stop Date	Action
1	GENERAL	N/A	N/A	Apply all general imputation rules to start and stop dates as indicated. Use the known actual dates and the imputed dates to derive each medication flag as described in the remaining specifications.
2	GENERAL	All cases	Missing	If the stop date is missing, assume the medication is ONGOING unless otherwise directed by the Medical Monitor for a particular medication. (For example, a medication which is used only during surgery and which has a missing stop date might be treated as started and stopped on the same day. The instruction would be applied to all instances in which this occurred, not just an individual case.)
3	GENERAL	All cases	Partial	Impute a partial stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown)
4	GENERAL	Missing	All cases	If the start date is missing, assume the medication was ongoing prior to first dose.
5	GENERAL	Partial	All cases	Impute a partial start date as earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown).
Prior medications are medications that were used before the first dose of study drug was administered.				

6	PRIOR	Known (Actual)
		or
		Partial (Imputed)
			Known (Actual)
			or
			Partial (Imputed)
				If (medication start date < date of first dose) then assign as prior='Yes'.
				If (medication start date = date of first dose and medication stop date = date of first dose) then assign as prior='No'.
				If (medication start date ≥ date of first dose) and (medication stop date is > date of first dose) then assign as prior='No'.

Rule	Rule Type	Start Date	Stop Date	Action
7	PRIOR	Known (Actual) or Partial (Imputed)	Missing	If (medication start date < date of first dose) then assign as prior='Yes'; If (medication start date ≥ date of first dose) then assign as prior='No'.
8	PRIOR	Missing	Known (Actual) or Partial (Imputed)	If missing (medication start date) then assign as prior='Yes'.
9	PRIOR:	Missing	Missing	If missing (medication start date) then assign as prior= 'Yes',
Concomitant medications are medications which were used during the period of treatment with study drug, including: • Medications that started prior to C1D1 and ended (or were ongoing) on or after C1D1 • Medications that started and ended on C1D1 • Medications that started on or after C1D1 and before a decision to discontinue treatment for AGS-16C3F subjects or before the last dose of study medication for axitinib subjects.				

10	CONCOMITANT	Known (Actual)
		or
		Partial (Imputed)
			Known (Actual)
			or
			Partial (Imputed)
				If (medication stop date < date of first dose) then assign as concomitant = 'No'.
				If (medication start date = date of first dose) and (medication stop date = date of first dose) then assign as concomitant ='Yes'.
				AGS: If (medication start date < date of decision to end treatment) then assign as concomitant ='Yes'.
				AGS: If (medication start date = date of decision to end treatment) and (medication stop date = date of decision to end treatment) then assign as concomitant ='Yes'.
				AGS: If (medication start date > date of decision to end treatment) then assign as concomitant = 'No'.
				AXI: If (medication start date < date of last dose) then assign as concomitant ='Yes'.
				AXI: If (medication start date = date of last dose) and (medication stop date = date of last dose) then assign as concomitant ='Yes'.
				AXI: If (medication start date > date of last dose) then assign as concomitant = 'No'.

Rule	Rule Type	Start Date	Stop Date	Action
11	CONCOMITANT	Known (Actual) or Partial (Imputed)	Missing	AGS: If medication start date < date of decision to end treatment then assign as concomitant = 'Yes'. AXI: If medication start date < date of last dose then assign as concomitant = 'Yes'.
12	CONCOMITANT	Missing	Known (Actual) or Partial (Imputed)	If (medication stop date < date of first dose), assign as concomitant='No'. If (medication stop date ≥ date of first dose), assign as concomitant='Yes'.
13	CONCOMITANT	Missing	Missing	Assign as concomitant='Yes',
Post-treatment (or "post") medications are medications that began on or after the date of a decision to discontinue study drug for AGS-16C3F subjects or after the last dose of study drug for axitinib subjects.				
13	POST	Known (Actual) or Partial (Imputed)	Known (Actual) or Partial (Imputed)	AGS: If (medication start date ≤ date of decision to end treatment) then post='No'. AGS: If (medication start date ≥ date of decision to end treatment) and (medication stop date > date of decision to end treatment) then post='Yes.' AXI: If (medication start date ≤ date of last dose) then post='No'. AXI: If (medication start date ≥ date of last dose) and (medication stop date > date of last dose) then post='Yes.'
14	POST	Known (Actual) or Partial (Imputed)	Missing	AGS: If (medication start date ≥ date of decision to stop treatment) then assign as post="Yes". AGS: If (medication start date < date of decision to stop treatment) then assign as post="No". AXI: If (medication start date ≥ date of last dose) then assign as post="Yes". AXI: If (medication start date < date of last dose) then assign as post="No".
15	POST	Missing	Known (Actual) or Partial (Imputed)	If the medication start date is missing, assign as post='No'.

Rule	Rule Type	Start Date	Stop Date	Action
16	POST	Missing	Missing	If the medication start date is missing, assign as post='No'.

APPENDIX 3. LIST OF EXCLUDED CONCOMITANT MEDICATIONS

The following medications are likely to influence evaluation of the safety, pharmacokinetics or efficacy of AGS-16C3F or axitinib, and will be strictly prohibited. If the investigator determines that any of the following medications are necessary to provide adequate medical support, the subject must be withdrawn from further study drug administration:

AGS-16C3F	Axitinib
Chemotherapy, radiotherapy, immunotherapy, monoclonal antibody therapy, or other medications intended for antitumor activity (i.e., other than AGS-16C3F) • Note: Palliative radiation is permitted provided it does not interfere with subject safety and efficacy of the study drug being administered. Palliative radiation cannot be administered to any target lesion.	Chemotherapy, radiotherapy, immunotherapy, monoclonal antibody therapy, or other medications intended for antitumor activity (i.e., other than axitinib) • Note: Palliative radiation is permitted provided it does not interfere with subject safety and efficacy of the study drug being administered. Palliative radiation cannot be administered to any target lesion.
Investigational products or therapy other than AGS-16C3F	Investigational products or therapy Strong CYP3A4/5 inducers[1]

[1] *The drugs defined as CYP Substrates with Narrow Therapeutic Range by FDA's guidance - Guidance for Industry Drug Interaction Studies - Study Design, Data analysis, and Implications for Dosing and Labeling): http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM292362.pdf)*

Examples of excluded inhibitors, inducers and substrates for enzymes and transporters are provided below. Subjects must have discontinued treatment with any of the following for at least 2 weeks prior to the first dose of study drug. This is not intended to be a comprehensive list. There could also be additional new drugs and marketed drugs that could be identified as inhibitors / inducers with continued research. Investigators will consult individual drug labels to determine liability of the drugs.

Strong CYP3A Inhibitors	
boceprevir	nefazodone
clarithromycin	nelfinavir
conivaptan	posaconazole
grapefruit juice	ritonavir
indinavir	saquinavir
itraconazole	telaprevir
ketoconazole	telithromycin
lopinavir/ritonavir	voriconazole
mibefradil	

CYP3A4/5 Inducers	
carbamazepine	rifapentin
dexamethasone	phenobarbital
rifabutin	phenytoin
rifampin	St. John's wort

Additional information for inhibitors/inducers/substrates of enzymes/transporters can be found in FDA's guidance (Guidance for Industry Drug Interaction Studies - Study Design, Data analysis, and Implications for Dosing and Labeling)[1] and from the Division of Clinical Pharmacology of Indiana University[2].

1. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM292362.pdf
2. http://medicine.iupui.edu/clinpharm/ddis/main-table/