CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

202324Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES-TEAM LEADER'S MEMO

NDA/BLA Serial Number:	202-324 / S-000
Drug Name:	Inlyta® (axitinib)
Indication(s):	Treatment of patients with advanced renal cell carcinoma.
Applicant:	Pfizer, Inc.
Date(s):	Submitted: April 14, 2011 PDUFA: February 14, 2012 Review Completed: January 11, 2012
Review Priority:	Standard
Biometrics Division:	Division of Biometrics V (HFD-711)
Primary Reviewer:	Somesh Chattopadhyay, Ph.D.
Secondary Reviewer:	Shenghui Tang, Ph.D., Team Leader
Concurring reviewer:	Rajeshwari Sridhara, Ph.D., Director
Medical Division:	Division of Oncology Products 1 (HFD-150)
Clinical Team:	Amy McKee, M.D., Medical Reviewer John R. Johnson, M.D., Medical Team Leader
Project Manager:	Ms. Lisa Skarupa

In the past six years, the treatment options for patients with advanced RCC have increased from IFN-α and IL-2 to six new agents with two different modes of actions: vascular endothelial growth factor receptor (VEGF-R) inhibitors sorafenib, sunitinib, and pazopanib and VEGF antibody bevacizumab; and mammalian target of rapamycin (mTOR) inhibitors temsirolimus and everolimus. All of the approvals for advanced RCC since 2005 have been given the broad indication of advanced RCC, except everolimus, which received a second-line indication. Most of the trials to support these broad indications were conducted in treatment-naïve patients; however, the pivotal trials for both sorafenib and pazopanib had mixed populations of treatment-naïve patients, patients who had received cytokine regimens, or patients who had received other regimens such as traditional chemotherapies or hormonal agents.

The applicant has submitted results from one multicenter, phase III, randomized, open-label, clinical trial (Study A4061032) comparing axitinib, a new molecular entity (NME), to sorafenib in patients with metastatic renal cell carcinoma (RCC) following failure of one prior systemic first line regimen containing one or more of the following: sunitinib, bevacizumab + IFN-α, temsirolimus, or cytokine(s). Study A4061032 randomized 723 patients in a 1:1 ratio to receive either axitinib at a starting dose of 5 mg twice daily or sorafenib 400 mg twice daily. The randomization was stratified by ECOG performance status and prior therapy. The primary endpoint was progression-free survival (PFS) based on the radiologic assessment by an independent review committee (IRC). The secondary endpoints included investigator-assessed PFS, overall survival (OS), objective response rate (ORR) as assessed by IRC, and duration of response. The axitinib arm showed statistically significant improvement over sorafenib in PFS as assessed by IRC in all randomized patients. The median PFS was 6.7 months in the axitinib arm and 4.7 months in the sorafenib arm with a hazard ratio (HR) of 0.67 (95% CI: 0.55-0.81). The difference in median PFS for patients previously treated with cytokines was 5.6 months (HR: 0.47; 95% CI 0.32-0.68), whereas the difference in patients previously treated with sunitinib was 1.4 months (HR: 0.74; 95% CI: 0.58-0.96). The study did not show difference in OS between axitinib and sorafenib arms (HR: 0.97; 95% CI: 0.80-1.17). For further details regarding the designs, data analyses, and results of Study A4061032, please refer to the statistical review by Dr. Somesh Chattopadhyay (January 11, 2012).

The application was discussed at the Oncologic Drug Advisory Committee meeting on December 7, 2011. The committee voted unanimously in favor of axitinib to the question whether the benefit-risk ratio of axitinib is favorable.

This team leader concurs with the recommendations and conclusions of the statistical reviewer (Dr. Somesh Chattopadhyay) of this application. The statistical results from Study A4061032 provide adequate evidence to support the PFS claim proposed in the NDA.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

SHENGHUI TANG
01/11/2012

RAJESHWARI SRIDHARA
01/11/2012

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Science
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA Serial Number:	202-324 / S-000
Drug Name:	Inlyta® (axitinib)
Indication(s):	Treatment of patients with advanced renal cell carcinoma.
Applicant:	Pfizer, Inc.
Date(s):	Submitted: April 14, 2011
	PDUFA: February 14, 2012
	Review Completed: January 9, 2012
Review Priority:	Standard
Biometrics Division:	Division of Biometrics V (HFD-711)
Statistical Reviewer:	Somesh Chattopadhyay, Ph.D.
Concurring Reviewers:	Shenghui Tang, Ph.D., Team Leader
	Rajeshwari Sridhara, Ph.D., Director
Medical Division:	Division of Oncology Products 1 (HFD-150)
Clinical Team:	Amy McKee, M.D., Medical Reviewer
	John R. Johnson, M.D., Medical Team Leader
Project Manager:	Ms. Lisa Skarupa

Keywords: Intent-to-treat, interim analysis, Kaplan-Meier product limit, logrank test, multiple endpoints, proportional hazards, randomization, stratification, subgroup analysis, survival analysis.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1. EXECUTIVE SUMMARY

The applicant has submitted results from one multicenter, phase III, randomized, open-label, clinical trial (Study A4061032) comparing axitinib, a new molecular entity (NME), to sorafenib in patients with metastatic renal cell carcinoma (RCC) following failure of one prior systemic first line regimen containing one or more of the following: sunitinib, bevacizumab + IFN-α, temsirolimus, or cytokine(s). The axitinib arm showed statistically significant improvement over sorafenib in progression-free survival (PFS) as assessed by independent radiology committee in all randomized patients. However, the axitinib arm did not show statistically significant improvement with respect to overall survival (OS). The application was discussed at the Oncologic Drug Advisory Committee meeting on December 7, 2011. The committee voted unanimously in favor of axitinib to the question whether the benefit-risk ratio of axitinib is favorable. The statistical results provide adequate evidence to support the PFS claim proposed in the NDA.

This application is based on one Phase III trial (Study A4061032) and three uncontrolled single-arm Phase II studies (A4061012, A4061035 and A4061023). This review is primarily based on the Phase III study. In Study A4061032 patients were randomized in a 1:1 ratio to receive axitinib at a starting dose of 5 mg twice daily orally with food or sorafenib at a starting dose of 400 mg twice daily orally without food. The randomization was stratified by ECOG performance status (0 vs. 1) and by prior therapy (sunitinib-containing regimens vs. bevacizumab-containing regimens vs. temsirolimus-containing regimens vs. cytokine-containing regimens). The study was initiated on September 15, 2008. The data cut-off date was August 31, 2010. A total of 723 patients were randomized, 361 to axitinib and 362 to sorafenib. Patients were enrolled at 175 centers in 22 countries. The primary efficacy endpoint was progression-free survival (PFS) as assessed by the independent radiology committee (IRC) review. The secondary efficacy endpoints were investigator-assessed PFS, overall survival (OS), objective response rate (ORR) as assessed by IRC review and duration of response.

The axitinib arm showed statistically significant improvement over sorafenib with respect to PFS as assessed by the IRC in the full analysis set (FAS) [hazard ratio=0.667, 95% confidence interval: (0.546, 0.814), log-rank test stratified by ECOG performance status and prior therapy, one-sided p-value<0.0001]. The median PFS and its 95% confidence interval in axitinib and sorafenib arms were 6.7 months [95% CI: (6.3 months, 8.4 months)] and 4.7 months [95% CI: (4.6 months, 5.6 months)], respectively. Analysis of the primary endpoint PFS as assessed by IRC is shown in Table 4 and the Kaplan Meier plot is shown in Figure 2. The study did not show a difference in OS between axitinib and sorafenib arms in the FAS [hazard ratio=0.969, 95% confidence interval: (0.800, 1.174), stratified log-rank test, one-sided p-value=0.3744]. Analysis of OS is shown in Table 7 and the Kaplan Meier plot of OS is presented in Figure 5. No adjustment to the level of significance was made for multiple secondary endpoints. Therefore p-values are not interpretable for the secondary endpoints. The objective response rate was 19.4% in the axitinib arm and 9.4% in the sorafenib arm based on the responses assessed by the IRC.

2. INTRODUCTION

2.1. Overview

Renal cell carcinoma (RCC) is the third leading urologic cancer. About thirty percent of patients with RCC have metastatic disease at the time of diagnosis, and a significant proportion of patients with localized disease treated with curative nephrectomy relapse subsequently with metastatic disease. The most frequent locations of metastases are the lungs, mediastinum, bone, liver, and brain. Metastatic RCC is currently incurable with a low 5-year survival rate and is frequently associated with a quality-of-life burden, based on physical, psychological, and social criteria.

Clear cell RCC, which represents 75% to 85% of the RCC population, frequently displays allelic loss on chromosome 3p, accompanied by mutational inactivation of the von Hippel- Lindau (VHL) tumor suppressor gene. VHL-associated RCC are known for their vascularity, and these tumors produce high levels of VEGF. In addition, recent studies suggest that in sporadic clear cell RCC, increased expression of VEGF is closely correlated with neovascularization, which is a prerequisite of tumor growth and metastasis. RCC tumors over express the receptors for these peptides. These ligands and receptors may be involved in the autocrine stimulation of tumor cell growth, or in the paracrine stimulation of neovascular or stromal fibroblast growth that supports tumor expansion. For these reasons, the angiogenic pathway is a logical target of potential therapies for advanced RCC.

2.1.1. Background

Angiogenesis, the formation and growth of new blood vessels, is a fundamental step in the transition of tumors from a dormant to a malignant state. The angiogenic process is initiated by various factors including vascular endothelial growth factor (VEGF), a cell-specific glycoprotein that plays a role in endothelial cell proliferation, survival, and migration by binding to tyrosine kinase receptors. These receptors are implicated in pathologic angiogenesis, tumor growth, and metastatic progression of cancer. The inhibition of angiogenesis through effects on VEGF has been a major target in the development of cancer therapies. Axitinib, a substituted indazole derivative, is an oral, potent, and selective tyrosine kinase inhibitor (TKI) of vascular endothelial growth factor receptor (VEGFR) 1, 2, and 3.

2.1.2. Regulatory History

Axitinib is a new molecular entity (NME). This application is based on a single randomized Phase III and three single arm Phase II trials. The trials were conducted under IND 63,662. The study protocol for the Phase III trial A4061032 was submitted for a Special Protocol Assessment (SPA) on December 18, 2007. FDA sent a non-agreement letter on January 31, 2008. The sponsor resubmitted the SPA on March 14, 2008 making modification to the protocol according to FDA's recommendations. FDA sent an agreement letter on April 17, 2008. The Pre-NDA

meeting was scheduled to be held on February 28, 2011. The sponsor cancelled the meeting after receiving FDA responses to their questions.

2.1.3. Specific Studies Reviewed

This application is based on one Phase III trial (Study A4061032) and three uncontrolled Phase II studies (A4061012, A4061035 and A4061023). Studies A4061012 and A4061035 were single-arm, single-agent Phase II trials in patients with cytokine-refractory advanced RCC. Study A4061023 was a single-arm, single-agent Phase II trial in patients with sorafenib-refractory advanced RCC.

This review is primarily based on the Phase III study. Study A4061032 was a multicenter, international, randomized, controlled, Phase III study to evaluate the efficacy of axitinib compared to sorafenib in patients with metastatic RCC following failure of one prior systemic first line regimen containing one or more of the following: sunitinib, bevacizumab + IFN-α, temsirolimus, or cytokine(s). Patients were randomized in a 1:1 ratio to receive axitinib at a starting dose of 5 mg twice daily orally with food or sorafenib at a starting dose of 400 mg twice daily orally without food. The randomization was stratified by ECOG performance status (0 vs. 1) and by prior therapy (sunitinib-containing regimens vs. bevacizumab-containing regimens vs. temsirolimus-containing regimens vs. cytokine-containing regimens). The study was initiated on September 15, 2008. The data cut-off date was August 31, 2010.

A total of 723 patients were randomized, 361 to axitinib and 362 to sorafenib. Of the randomized patients, 523 were men and 200 were women, 547 were White, and the median age was 61 years (age range: 20 to 82 years). Randomized patients were enrolled at 175 centers in 22 countries. There were 169 patients from US.

2.2. Data Sources

Data used for this review are from the electronic submissions dated April 14, 2011, June 9, 2011 and December 14, 2011. The paths are \\Cdsesub1\EVSPROD\NDA202324\0000\m5\datasets\a4061032, \\Cdsesub1\EVSPROD\NDA202324\0004\m5\datasets\a4061032 and \\Cdsesub1\EVSPROD\NDA202324\0022\m5\datasets\a4061032, respectively.

3. STATISTICAL EVALUATION

3.1. Data and Analysis Quality

Overall the data and analysis quality of the submission was acceptable for the reviewer to be able to perform the statistical review. However there were some problems with the datasets in the original and subsequent submissions.

The submission had the following problems.

- Some variables were missing from the datasets in the original submission.
- In the original submission, parts of the data definition file were unclear including the lack of specifications of the key variables for the datasets with multiple records per patient and unclear variable labels.

The applicant corrected the above problems and resubmitted new datasets and data definition files to address the above problems.

3.2. Evaluation of Efficacy

The applicant has submitted efficacy results from one Phase III study (Study A4061032) titled "Axitinib (AG-013736) as second-line therapy for metastatic renal cell cancer: Axis Trial" and three uncontrolled Phase II studies (A4061012, A4061035 and A4061023). Studies A4061012 and A4061035 were single-arm, single-agent Phase II trials in patients with cytokine-refractory advanced RCC. Study A4061023 was a single-arm, single-agent Phase II trial in patients with sorafenib-refractory advanced RCC. This review will be primarily based on the Phase III study.

3.2.1. Study Objectives

3.2.1.1. Primary Objective

The primary objective of Study A4061032 was to compare the progression-free survival (PFS) of patients with mRCC receiving axitinib versus sorafenib following failure of one prior systemic first-line regimen containing one or more of the following: sunitinib, bevacizumab + IFN-α, temsirolimus, or cytokine(s).

3.2.1.2. Secondary Objectives

The secondary objectives were to:

- Compare the overall survival (OS) of patients between two arms;
- Compare the objective response rate (ORR) of patients between two arms;
- Evaluate the safety and tolerability of axitinib;
- Estimate the duration of response (DR) of patients in each arm; and

- Compare the kidney-specific symptoms and health status of patients between two arms, as measured by the Functional Assessment of Cancer Therapy Kidney Symptom Index (FKSI) and European Quality of Life (EuroQol) EQ-5D self-report questionnaire (EQ-5D).

3.2.2. Study Design

This study was a multicenter, international, open-label, randomized, controlled, Phase III study to evaluate the efficacy of axitinib compared to sorafenib in patients with metastatic RCC following failure of 1 prior systemic first line regimen containing one or more of the following: sunitinib, bevacizumab + IFN-α, temsirolimus, or cytokine(s).

Patients were randomized in a 1:1 ratio to receive axitinib at a starting dose of 5 mg twice daily orally with food or sorafenib at a starting dose of 400 mg twice daily orally without food (at least 1 hour before or 2 hours after eating). A centralized registration and randomization system (Interactive Voice Response System [IVRS]) was used for randomization. The patients were stratified by ECOG performance status (0 vs. 1) and by prior therapy (sunitinib-containing regimens vs. bevacizumab-containing regimens vs. temsirolimus-containing regimens vs. cytokine-containing regimens).

Study treatment was to continue until disease progression, intolerable adverse drug reactions, or withdrawal of consent. Study treatment was required to begin within 7 days of randomization.

Patients who discontinued treatment on this study could receive subsequent therapy based on the judgment of the treating physician. Patients were not offered axitinib or sorafenib as a subsequent therapy while on-study.

Major inclusion criteria included histologically or cytologically confirmed mRCC with a component of clear cell subtype, evidence of unidimensionally measurable disease (i.e., ≥1 malignant tumor mass that could have been accurately measured in at least 1 dimension ≥ 20 mm with conventional computed tomography [CT] scan or magnetic resonance imaging [MRI] scan, or ≥ 10 mm with spiral CT scan using a 5 mm or smaller contiguous reconstruction algorithm), progressive disease per RECIST 1.0 after 1 prior systemic first-line regimen for mRCC containing one or more of sunitinib, bevacizumab + IFN- α, temsirolimus and cytokine(s), adequate organ function and ECOG performance status 0 or 1.

3.2.3. Schedule of Assessments

Baseline tumor assessments required CT/MRI (no chest x-ray) of the chest, abdomen, and pelvis and a bone scan at the minimum within 28 days prior to randomization. CT or MRI of brain was required at baseline. Subsequent CT or MRI of the brain could have been performed if clinically indicated, as determined by the treating physician. On-study tumor assessments were to be performed every 6 weeks for the first 12 weeks, and then every 8 weeks by calendar to determine PFS. If a baseline bone scan showed metastatic lesions, the lesions were confirmed with

concomitant x-ray, CT, MRI, and bone scans and bone imaging was required at the time points matched with the CT/MRI evaluations. All scans were sent to the independent review committee (IRC). All patients were evaluated for response according to RECIST (Version 1.0)

Patients removed from axitinib treatment for intolerable toxicity were to be followed with regular tumor assessments until disease progression or the start of new treatment, and for survival thereafter. Patients removed from sorafenib treatment for intolerable toxicity were to be followed for survival.

3.2.4. Efficacy Endpoints

Primary endpoints:

- Progression-free survival (PFS) as assessed by the IRC.

Secondary endpoints:

- Progression-free survival (PFS) as assessed by the investigator.
- Overall survival (OS)
- Objective response rate (ORR)
- Duration of response (DR)

Patient reported outcome endpoint:

- Kidney-specific symptoms as measured by Functional Assessment of Cancer Therapy Kidney Symptom Index (FKSI)
- Health status as measured by European Quality of Life (EuroQol) EQ-5D self-report questionnaire (EQ-5D)

PFS was defined as the time from randomization to first documentation of objective tumor progression or to death due to any cause, whichever occurred first. If tumor progression data included more than 1 date, the first date was used. For the purposes of endpoint definitions, the term 'on-study' includes the period from randomization until 28 days after the last dose of study medication.

PFS data were censored on the date of the last tumor assessment (on-study) documenting absence of PD for patients

- Who were alive, on-study, and progression free at the time of the analysis;
- Who had at least 1, on-study disease assessment and discontinued treatment without documented disease progression and without death on-study;
- For whom documentation of disease progression or death occurred after ≥2 consecutive missed tumor assessments; or
- Who were given antitumor treatment, other than the study treatment, before documented disease progression.

Patients lacking an evaluation of their disease at baseline had their event time censored on the date of randomization. Patients lacking an evaluation of tumor response after randomization also

had their event time censored on the date of randomization unless death occurred prior to the first planned assessment (in which case the death was an event).

OS was defined as the time from the date of randomization to the date of death due to any cause. For patients still alive at the time of the analysis, the OS time was censored on the last date they were known to be alive. Patients lacking data beyond randomization had their OS times censored at the date of randomization.

ORR was defined as the percent of patients with confirmed CR or confirmed PR (by repeat imaging study at least 4 weeks after initial documentation of response) according to RECIST criteria, relative to all randomized patients. Third-party blinded review and qualification were performed retrospectively by the IRC. Patients who did not have on-study radiographic tumor re-evaluation or who died, progressed, or dropped out for any reason before reaching a CR or PR were counted as nonresponders in the assessment of ORR. A patient who initially met the criteria for a PR and then subsequently became a confirmed CR was assigned a best response of CR.

DR was defined as the time from the first documentation of objective tumor response (CR or PR), that was subsequently confirmed, to the first documentation of PD or to death due to any cause, whichever occurred first. If tumor progression data included more than 1 date, the first date was used. Patients who achieved a PR and then a CR had times calculated using the date of the PR as the first day. DR was only calculated for the subgroup of patients with a confirmed objective tumor response.

DR data were censored on the date of the last tumor assessment documenting absence of progressive disease for patients

- Who were alive, on-study, and progression free at the time of the analysis;
- Who discontinued treatment without documented disease progression and without death on-study;
- For whom documentation of PD or death occurred after ≥2 consecutive missed tumor assessments; or
- Who were given antitumor treatment, other than the study treatment, prior to documented disease progression.

The FKSI contains 15 questions and includes a 9-question subscale known as the FKSI-Disease Related Symptoms (FKSI-DRS) subscale that specifically measures symptoms related to advanced kidney cancer disease. This subscale includes the following 9 items: lack of energy, pain, losing weight, bone pain, fatigue, short of breath, coughing, bothered by fevers, and hematuria. Each question in the FKSI questionnaire is answered on a 5-point Likert-type scale ranging from 0 to 4 (0=not at all, 1=a little bit, 2=somewhat, 3=quite a bit, 4=very much). For some questions, the answers are the item scores; for other questions, the answers are reverse coded to create the item scores. The total FKSI score is the sum of the 15 individual item scores. The total FKSI-DRS score is the sum of its 9 individual item scores.

The EQ-5D is a brief, self-administered generic health status instrument consisting of 2 parts. In the first part, respondents are asked to describe their current health state on 5 dimensions

(mobility, self-care, usual activities, pain or discomfort, and anxiety or depression) with each dimension having 3 levels of function (1=no problem, 2=some problem, and 3=extreme problem). The second part is a patient's self-rating of current health state on a visual analog scale (EQ-VAS) with endpoints labeled 'best imaginable health state' and 'worst imaginable health state'. The EQ-5D endpoints are the EQ-5D Index, which is derived by combining 1 level from each of the 5 dimensions and converting it to a single summary index or health utility value, and the EQ-VAS score, which ranges from 0 for worst imaginable health state to 100 for best imaginable health state. Overall, scores for the EQ-5D index range from -0.594 to 1, with low scores representing a higher level of dysfunction.

Reviewer's comment:

Patient-reported outcome endpoints are subject to bias in an open-label study. Therefore, this review considers these endpoints in this open-label trial as exploratory.

3.2.5. Sample Size Considerations

The sample size for this study was calculated based on PFS. In order to detect a median PFS of 7 months in the axitinib arm versus a median PFS of 5 months in the sorafenib arm at a one-sided significance level of 0.025 with 90% power, 1:1 randomization and a futility analysis at 50% PFS events, a total of 409 PFS events were required. Assuming a planned accrual period of 18 months and a follow-up period of approximately 5 months, it was estimated by the applicant that approximately 650 patients would need to be enrolled in order to observe 409 PFS events.

The sample size described above also allowed to adequately power the study for OS. In order to detect a median OS of 23.7 months in the axitinib arm versus a median OS of 18 months in the sorafenib arm at a one-sided significance level of 0.025 with 80% power, 1:1 randomization and two interim analyses, one after approximately 20% events and the other after approximately 46% events, a total of 417 OS events were required. The sample size 650 would also be sufficient to observe 417 OS events with approximately 37 months follow-up.

Other secondary and supportive analyses were tested at a significance level of 0.025 (1-sided test). No adjustments were planned for multiple testing/comparisons in those secondary and supportive hypothesis tests.

Reviewer's comments:

1. The initial target enrollment was 540 patients. The applicant (then sponsor) increased the sample size to 650 in the Protocol Amendment 4 dated November 16, 2009 and the corresponding Statistical Analysis Plan (version 1.2) dated December 7, 2009. The applicant claims that this change was made due to an underestimation of the dropout rate in the original protocol. However, Table 1 shows that only 21 patients in the axitinib arm (5.8% of the all axitinib patients) and 37 patients in the sorafenib arm (10.2% of the all sorafenib patients) were censored prior to data cut-off. However, changing the sample size alone without any change in the number of events does not affect the error rates.

Table 1: Reasons for Censoring of IRC-assessed PFS

Reason for PFS (by IRC) Censoring	Axitinib (169 censored)		Sorafenib (152 ensored)	
No baseline on study assessments	14	(8.3%)	28	(18.4%)
Alive, on-study and progression-free at the time of the analysis	148	(87.6%)	115	(75.7%)
At least 1 on-study assessment and discontinued treatment prior to documented PD on-study	4	(2.4%)	4	(2.6%)
PD or death occurred after ≥ 2 consecutive, missed assessments	1	(0.6%)	3	(2.0%)
PD occurred after given new anti-tumor treatment	2	(1.2%)	0	(0%)
Withdrew consent for follow-up	0	(0%)	0	(0%)
Lost to follow-up	0	(0%)	2	(1.3%)

2. The actual sample size for the study was 723, which is even much higher than 650, the increased sample size that was determined at the time of Protocol Amendment 4. There was no justification provided for such a big difference between the planned and actual sample size. According to the reviewer, it appears that the fast accrual is a possible reason for accruing many more patients than targeted. The number of PFS events as assessed by IRC was 402 at the final analysis which is very close to the targeted number of events 409. Also, it should be noted that the sample size increase was not because of the guideline for sample size adjustment based on the conditional power as planned in the protocol.

3. The interim analysis for futility was planned to be performed after approximately 204 PFS events (approximately 50% of the total number of events as assessed by the IRC). The actual interim analysis was performed with 289 PFS events (70.7% of the original 409 required events) observed. The applicant claims that because the interim analysis occurring later than planned, the required number of PFS events would need to be increased to 423 in order to maintain 90% power with the same design parameters. The applicant did not amend the protocol or the Statistical Analysis Plan to include this modification. Applicant's claim about the need to increase the number of events was not correct. Based on the reviewer's calculation using East version 5, the number of events should not be more than 414. The number of PFS events at the final analysis was 428 based on investigator's assessment and 402 based on IRC. The independent DMC met on June 11, 2010 to analyze the interim results; the DMC concluded that there were no safety concerns, the PFS results did not cross the futility boundary, no adjustment of sample size was necessary, and that the study should continue as planned.

3.2.6. Interim Analyses

The study was designed to have 1 interim analysis and the final analysis for PFS. The interim analysis and final analysis were based on the primary endpoint of PFS as assessed by the IRC. A formal efficacy boundary for rejecting the null hypothesis and futility boundary (nonbinding) for rejecting the alternative hypothesis were constructed by using the spending function methodology of Lan and DeMets in an asymmetric fashion. An O'Brien-Fleming stopping boundary was used for the efficacy and a Pampallona-Tsiatis Power boundary (with parameter 0.3) was used for the futility. Both type 1 error and type 2 error were preserved. Nonbinding for the futility implied that the futility boundary was constructed in such a way that it could be overruled if desired by the DMC without inflating the type 1 error and without decreasing the power.

The purposes for the interim analysis were to:

- Assess the safety of the study treatment;
- Stop the study early due to futility, or
- Adjust the sample size.

The interim analysis of efficacy and safety was planned to be performed after approximately 204 PFS events (approximately 50% of the total number of required events as assessed by the IRC). If the results of the interim analysis indicated serious safety concerns, it was planned that the Sponsor would consult with regulatory authorities regarding stopping the clinical study. If exactly 204 expected events occurred at the time of the interim analysis, then the significance level for futility would be 0.2293.

To protect the integrity of the study and to preserve the type 1 error, a fraction of alpha was spent at the interim analysis (although there were no plans to stop the study early for efficacy based on PFS data at this interim analysis) based on an O'Brien-Fleming stopping boundary.

An interim analysis of OS was also planned at the time of the final PFS analysis. The purpose was to correlate OS outcome with PFS outcome. The nominal significance level for the interim and final analyses for OS were determined by using the Lan-DeMets procedure with an O'Brien-Fleming stopping rule. The overall nominal significance level for the efficacy analysis of OS was preserved at 0.025 (1-sided test).

According to the Lan-DeMets alpha spending function for O'Brien-Fleming boundary and the actual number of events, the alpha for the interim OS analysis with cut-off date August 31, 2010 and final analysis with the cut-off date November 1, 2011 would be 0.002 and 0.0244, respectively (calculated using East Version 5).

The sample size of the study could be adjusted using the method outlined by Cui, Hung and Wang (1999). After examining the data available at the interim analysis, the DMC statistician may have suggested an adaptive change in sample size, based on the following algorithm:

- If the conditional power at the interim analysis was ≥90%, no sample size adjustment was to be made;

- If the conditional power at the interim was ≥83%, but <90%, then the maximum number of events was to be increased, which may have resulted in a sample size increase. The total number of events for the study was not to be increased by more than 80 events; and
- If the conditional power at the interim was <83%, no sample size adjustment was to be made.

A sample size adjustment would have allowed the study to detect a HR of 0.769 (~30% increase in median PFS for axitinib) with at least 90% power. The overall type I error would have been preserved after the sample size adjustment.

The interim analysis for futility was planned to be performed after approximately 204 PFS events (approximately 50% of the total number of events as assessed by the IRC). The actual interim analysis was performed with 289 PFS events (70.7% of the original 409 required events) observed. The applicant claims that because the interim occurring later than planned, the required number of events would need to be increased to 423 in order to maintain 90% power with the same design parameters.

The independent DMC met on June 11, 2010 to analyze the interim results; the DMC concluded that there were no safety concerns, the PFS results did not cross the futility boundary, no adjustment of sample size was necessary, and that the study should continue as planned.

Reviewer's comments:

The interim analysis report that the DMC reviewed was not submitted with this application. Also refer to the reviewer's comments in Section 3.2.5 about the effect of the timing of the interim analysis on increasing the number of PFS events.

3.2.7. Efficacy Analysis Methods

3.2.7.1. Analysis Populations

The following analysis populations were used in this study:

Full Analysis Set (FAS) - All patients who were randomized, with study drug assignment designated according to initial randomization, regardless of whether patients received study drug or received a different drug from that to which they were randomized. The FAS was the primary population for evaluating all efficacy endpoints, as well as patient characteristics.

Safety Analysis Set (SA) - All patients who received at least 1 dose of study medication, with treatment assignments designated according to actual study treatment received. This SA set was the primary population for evaluating treatment administration/compliance and safety.

3.2.7.2. Analysis of Primary Endpoint

PFS, based on IRC assessment, was the primary efficacy endpoint. PFS was summarized for the FAS using Kaplan-Meier methods. The median event time for each treatment arm and corresponding 2-sided 95% CI for the median were provided for PFS. A one-sided log-rank test at α=0.025 stratified by ECOG PS and prior therapy was used to compare PFS between the 2 treatment arms. The hazard ratio and its 95% CI were estimated using a stratified Cox model. An unstratified one-sided log-rank test at α=0.025 and Cox regression model were used as secondary analyses for PFS.

Cox regression models were used to explore the potential influences of the stratification factors baseline patient characteristics (e.g., age, ethnic origin, sex, geographic region, MSKCC risk group) on the primary PFS endpoint. For each treatment arm, the median PFS and a 2-sided 95% CI were provided for each level of the stratification variables.

Preplanned analyses of the primary efficacy endpoint and all sensitivity analyses were performed for the stratification factors ECOG performance status (0 and 1) and prior treatment regimen (sunitinib-containing regimen, bevacizumab-containing regimen, temsirolimus-containing regimen, and cytokine-containing regimen). In addition, preplanned subgroup analyses were performed on the primary efficacy endpoint for the baseline patient characteristics of age (<65 years, ≥65 years), sex (male, female), ethnic origin (white, nonwhite), geographic region (Asia, Europe, North America, Other) and MSKCC risk group (favorable, intermediate, poor).

Reviewer's comment:

Although the subgroup analyses by the stratification factors (ECOG performance status and prior therapy) were pre-specified, there was no adjustment made in the Type I error rate for multiple subgroup analyses. Therefore, these subgroup analyses are considered exploratory by the reviewer.

3.2.7.3. Analysis of Secondary Endpoints

PFS, based on investigator assessment, was a secondary efficacy endpoint. A one-sided stratified log-rank test at α=0.025 was used to evaluate investigator assessed PFS, in the SA set (i.e., all patients who received at least one dose of study medication).

The number and percent of patients achieving objective response (CR or PR) were summarized along with the corresponding exact 2-sided 95% CI calculated using a method based on the F distribution. A Pearson χ^2 test (unstratified) and Cochran-Mantel-Haenszel test stratified by baseline stratification factors was used to compare ORR between the 2 treatment arms. For the unstratified analyses, point estimates of the rates for each treatment arm and difference of the rates between treatment arms were provided, along with the corresponding 2-sided 95% CIs, using an exact method based on the F distribution and using a normal approximation for constructing a CI for differences, respectively. For the stratified analyses, the relative risk ratio

estimator was used to contrast the treatment effects on the endpoint. Both a point estimate and a 2-sided 95% CI were calculated using a normal approximation.

Time-to-event endpoints, including OS and DR, were summarized using Kaplan-Meier methods and displayed graphically. DR was calculated for the subgroup of patients with objective disease response. Unstratified and stratified log-rank tests (1-sided, α=0.025) were used to compare OS between the 2 treatment arms. The median event time and 2-sided 95% CI for the median were provided for each endpoint. The hazard ratio and its 95% CI were estimated for OS. Additionally for each treatment arm, the median OS and a 2-sided 95% CI were provided for each level of the stratification variables.

Preplanned analyses of secondary efficacy endpoints, including OS, ORR and DR, were performed for the stratification factors ECOG performance status (0 and 1) and prior treatment regimen (sunitinib-containing regimen, bevacizumab-containing regimen, temsirolimus-containing regimen, and cytokine-containing regimen).

Reviewer's comment:

Type I error rate has not been adjusted for analysis of multiple secondary endpoints. Therefore, p-values for the secondary endpoints are not interpretable.

3.2.7.4. Patient Reported Outcome Analyses

FKSI-15 was the sum of the scores from the 15 FKSI questions. FKSI-15 was summarized using means (with standard deviations) and medians at each assessment point, based on the observed values as well as changes from baseline both within group and between groups (with 95% CIs for mean changes). Comparisons of the 2 treatments were based on a repeated measures analysis using a mixed effects model. The variables in the model were treatment, time, and treatment-by-time, with baseline as covariate and time assumed linear. The minimally important difference (MID) for FKSI was 3 to 5 points. FKSI-DRS was measured as the sum of the scores from the 9 FKSI-DRS questions. Analyses of FKSI-DRS, EQ-5D and EQ-VAS followed the same methodology as for FKSI-15. For a clinically meaningful difference, the MID for FKSI-DRS was assumed to be 2 to 3 points.

Reviewer's comment:

Patient-reported outcome endpoints are subject to bias in an open-label study. Therefore, this review considers those endpoints in this open-label trial exploratory.

3.2.8. Sponsor's Results and FDA Statistical Reviewer's Findings/Comments

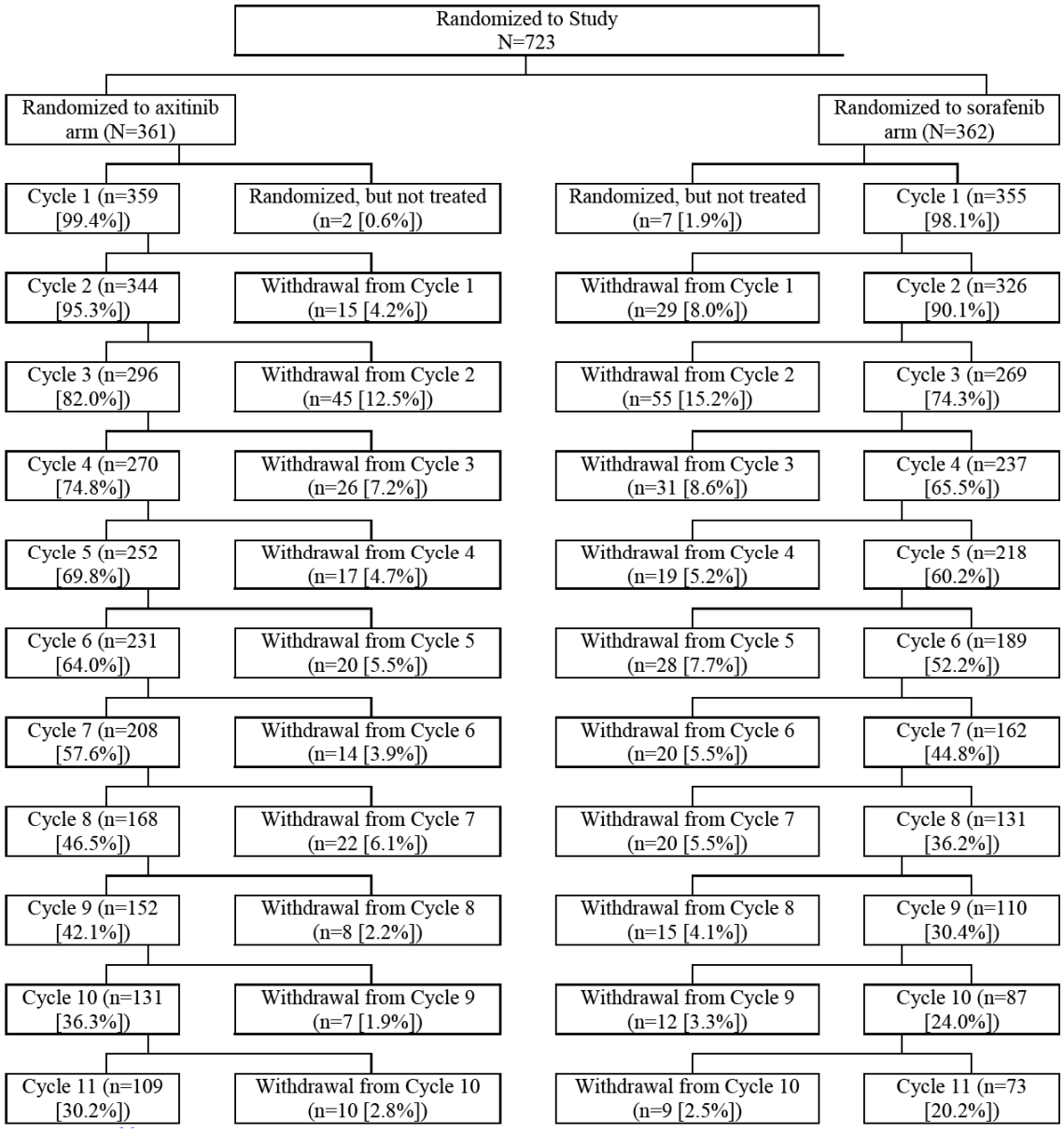
A total of 723 patients were randomized, 361 to receive either axitinib of whom 359 received treatment and 362 to sorafenib of whom 355 received treatment. Patients were recruited at 175

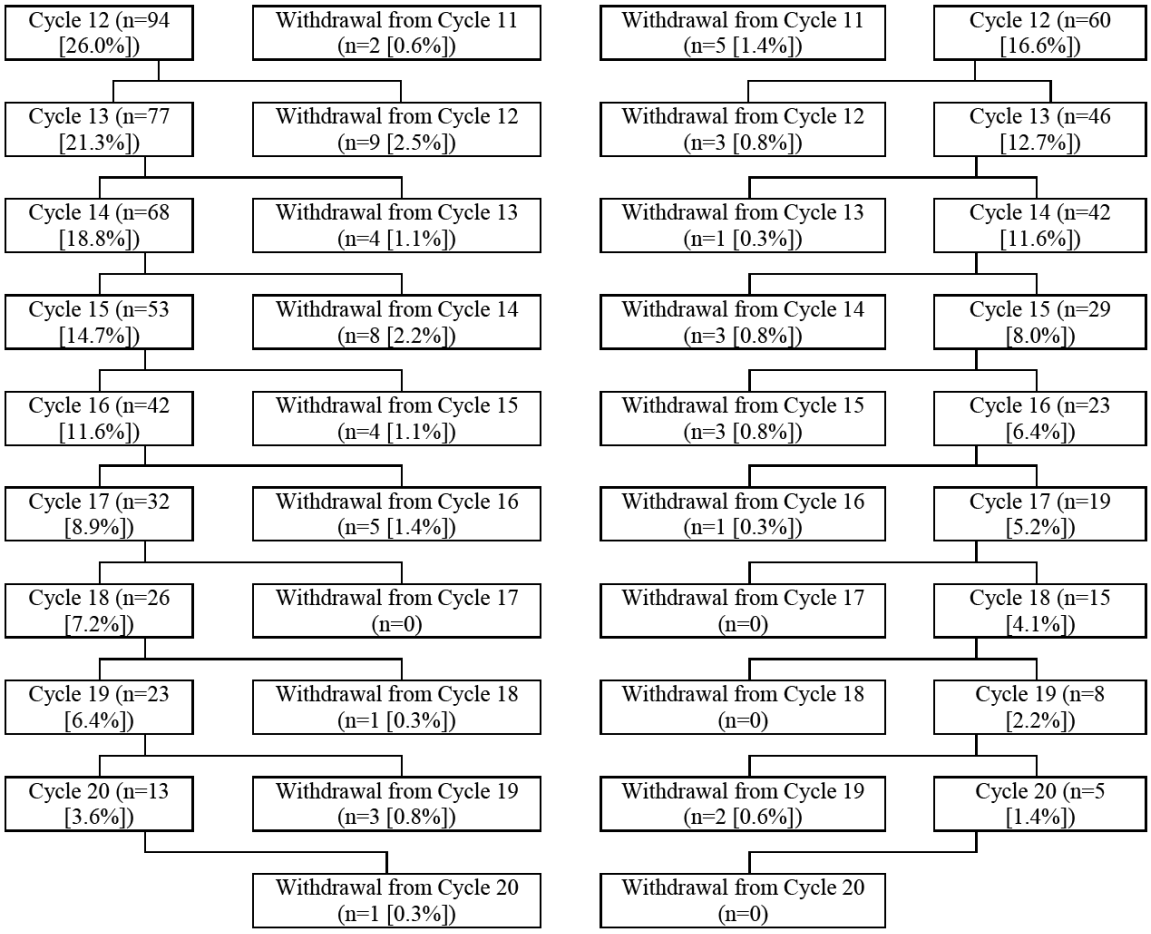
centers in 22 countries from September 15, 2008 to July 23, 2010. The data cut-off date for this study was August 31, 2010.

3.2.8.1. Patient Disposition

The patient disposition is presented in Figure 1.

Figure 1: Patient Disposition





Source: Clinical study report submitted in the NDA.

3.2.8.2. Baseline Characteristics

The treatment arms were well balanced with respect to general demographic characteristics (gender, race and age) at baseline. The study population consisted of 72.34% male. Majority of the patients were White (75.66%). The mean and median age of patients at randomization was 59.8 and 61 years, respectively, with an overall age range of 20 to 82 years. Approximately 34% of patients were elderly. The study recruited patients from 22 countries. Approximately 23% patients were from US. There was higher percentage of US patients in the sorafenib arm (25%) than in the axitinib arm (21%). A summary of demographic characteristics at baseline is presented in Table 2.

Stratification factors as baseline characteristics are presented in Table 3. Approximately 55% patients enrolled in this study had ECOG PS of 0, 45% had ECOG PS 1. Approximately 54% patients were previously treated with sunitinib, 35% with cytokines, 8% with bevacizumab+IFN-α and only 3% with temsirolimus.

Table 2: Demographic Characteristics: Gender, Race and Age at Randomization in FAS

		Axitinib (N=361)	Sorafenib (N=362)	All (N=723)
Gender	Female	96 (26.59%)	104 (28.73%)	200 (27.66%)
	Male	265 (73.41%)	258 (71.27%)	523 (72.34%)
Race	White	278 (77.01%)	269 (74.31%)	547 (75.66%)
	Black	1 (0.28%)	4 (1.10%)	5 (0.69%)
	Asian	77 (21.33%)	81 (22.38%)	158 (21.85%)
	Other	5 (1.39%)	8 (2.21%)	13 (1.80%)
Age Group in Years	<65	238 (65.93%)	238 (65.75%)	476 (65.84%)
	≥65	123 (34.07%)	124 (34.25%)	247 (34.16%)
Age in Years at Randomization	Mean, SD	59.71, 10.55	59.98, 10.14	59.84, 10.34
	Min, Max	20, 82	22, 80	20, 82
	Q1, Median, Q3	53, 61, 67	54, 61, 67	54, 61, 67
Geographic region	US	77 (21.33%)	92 (25.41%)	169 (23.37%)
	Non-US	284 (78.67%)	270 (74.59%)	554 (76.63%)

Table 3: Baseline Characteristics (Stratification Factors)

		Axitinib (N=361)	Sorafenib (N=362)	All (N=723)
Prior Therapy	Bevacizumab +IFN-α	29 (8.03%)	30 (8.29%)	59 (8.16%)
	Cytokine(s)	126 (34.90%)	125 (34.53%)	251 (34.72%)
	Sunitinib	194 (53.74%)	195 (53.87%)	389 (53.80%)
	Temsirolimus	12 (3.32%)	12 (3.31%)	24 (3.32%)
ECOG Performance Status	0	197 (54.57%)	199 (54.97%)	396 (54.77%)
	1	164 (45.43%)	163 (45.03%)	327 (45.23%)

3.2.8.3. Primary Efficacy Analysis

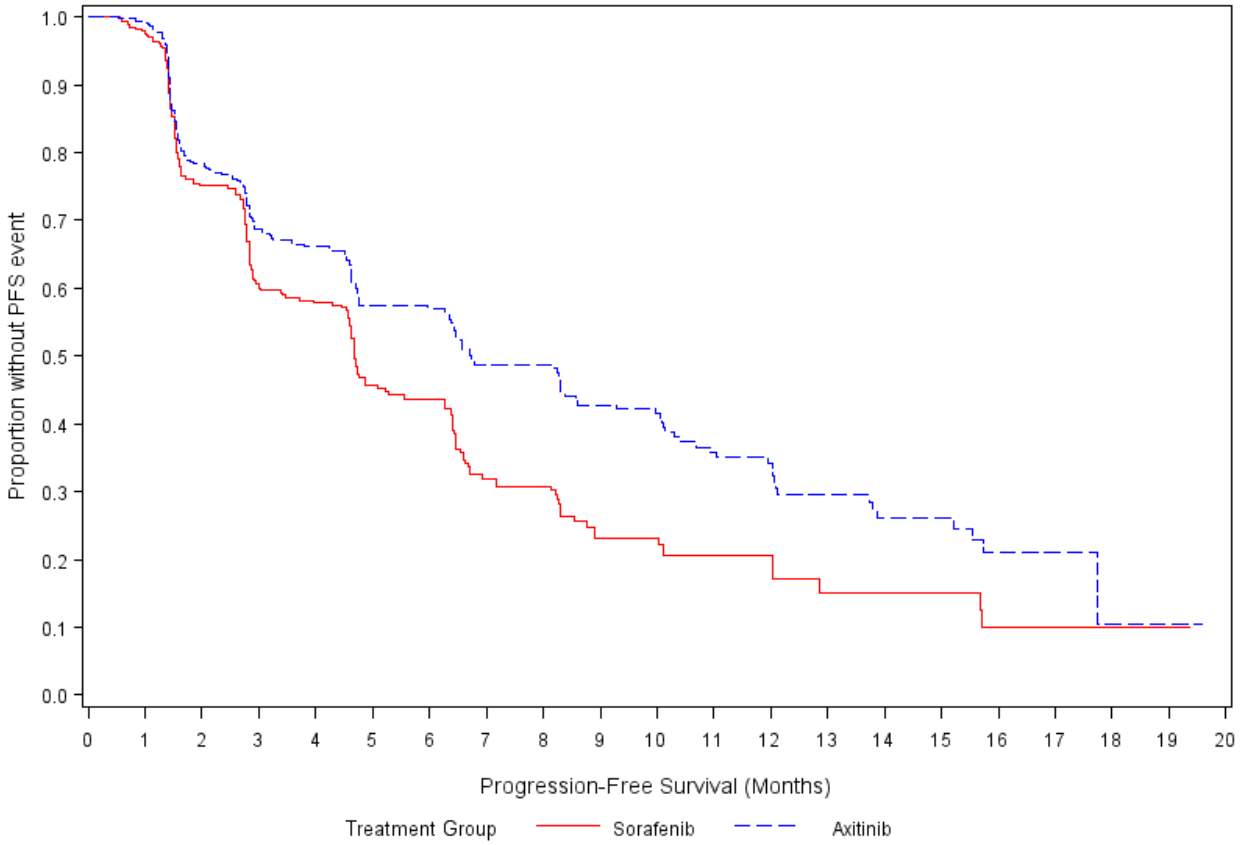
The primary efficacy analysis comparing progression-free survival (PFS) between axitinib and sorafenib in FAS based on independent radiologic committee review using log-rank test stratified by ECOG performance status and prior therapy, the same stratification factors used at randomization, is presented in Table 4. The corresponding Kaplan-Meier plots is given in Figure 2. The PFS improvement in axitinib arm over sorafenib arm was statistically significant (stratified log-rank test, nominal two-sided p-value < 0.0001). The PFS hazard ratios of axitinib over sorafenib using a stratified Cox model with the same stratification factors ECOG performance status and prior therapy and its 95% confidence intervals were 0.667 [95% CI: (0.546, 0.814)]. The difference in median PFS between two arms is approximately 2 months.

Table 4: Analysis of PFS Based on Independent Review in FAS

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)	P-value[3]
Axitinib	361	192 (53.19%)	6.7 (6.3, 8.4)	0.667 (0.546, 0.814)	<0.0001
Sorafenib	362	210 (58.01%)	4.7 (4.6, 5.6)		

[1]: Kaplan-Meier estimate. [2]: Based on stratified Cox model. [3]: Based on one-sided log-rank test, stratified by ECOG PS and prior therapy.

Figure 2: Kaplan-Meier Plot of PFS in FAS Based on Independent Review



Reviewer's Comment:

Although the PFS improvement in axitinib arm over the sorafenib arm is small in terms of the difference in median, sorafenib is an active control. Sorafenib's effect in this population is not exactly known; it is likely to be better than a placebo. Sorafenib is an approved drug for the treatment of mRCC.

3.2.8.4. Secondary Efficacy Analyses

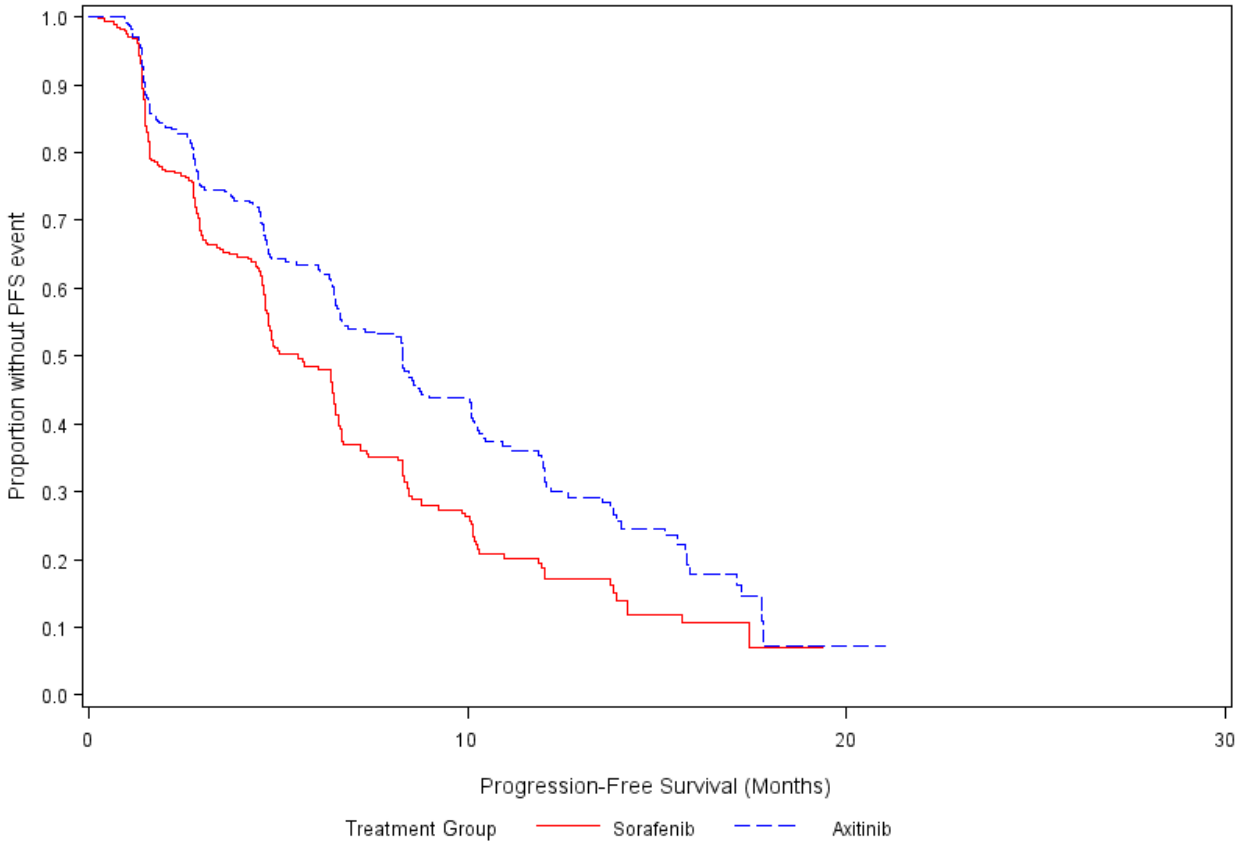
Investigator assessed PFS was a secondary endpoint. The analysis of PFS based on investigator assessment is presented in Table 5 and the Kaplan-Meier plot is presented in Figure 3. The hazard ratio and p-values are similar to those for PFS based on independent review. However the medians are higher than those assessed by independent review.

Table 5: Analysis of PFS Based on Investigator's Assessment in FAS

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)	P-value[3]
Axitinib	361	201 (55.68%)	8.3 (6.6, 9.0)	0.659 (0.543, 0.799)	<0.0001
Sorafenib	362	227 (62.71%)	5.6 (4.7, 6.5)		

[1]: Kaplan-Meier estimate. [2]: Based on stratified Cox model. [3]: Based on one-sided log-rank test, stratified by ECOG PS and prior therapy.

Figure 3: Kaplan-Meier Plot of PFS Based on Investigator's Assessment in FAS



At the time of the original NDA submission, the applicant submitted the results of the interim OS analysis based on data cut-off of 31 August, 2010. On December 14, 2011, the applicant submitted the mature OS data with 425 events based on a cut-off date of November 1, 2011. The analyses of OS in FAS based on the data with cut-off dates August 31, 2010 (interim analysis) and November 1, 2011 (final analysis) are presented in Table 6 and Table 7, respectively. The corresponding Kaplan-Meier plots are given in Figure 4 and Figure 5, respectively. There was no statistically significant difference between axitinib and sorafenib arms with respect to OS at both interim and final analyses (log-rank test stratified by ECOG performance status and prior therapy, nominal one-sided p-value 0.5253 at the interim analysis and 0.3744 at the final analysis). The Kaplan-Meier plots show that OS in the two arms are very similar. The hazard ratio for OS was 1.009 [95% CI: (0.774, 1.313)] at the interim analysis and 0.969 [95% CI: (0.800, 1,174)] at the final analysis.

Table 6: Analysis of OS in FAS (Interim Analysis, August 31, 2010 Cut-off)

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)	P-value[3]
Axitinib	361	113 (31.30%)	NE (15.9, NE)	1.009 (0.774, 1.313)	0.5253
Sorafenib	362	110 (30.39%)	18.9 (18.0, NE)		

[1]: Kaplan-Meier estimate. [2]: Based on stratified Cox model. [3]: Based on one-sided log-rank test, stratified by ECOG PS and prior therapy.

Figure 4: Kaplan-Meier Plot of OS in FAS (Interim Analysis, August 31, 2010 Cut-off)

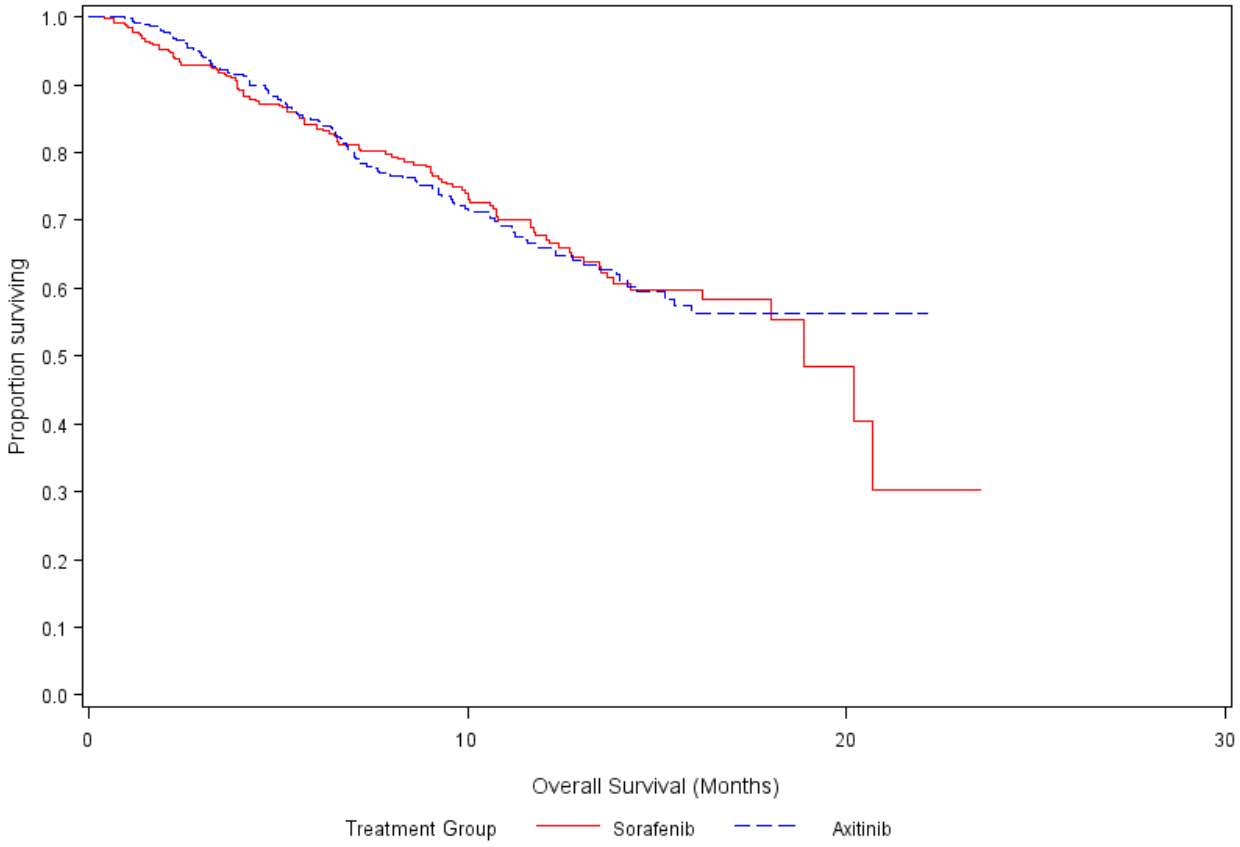
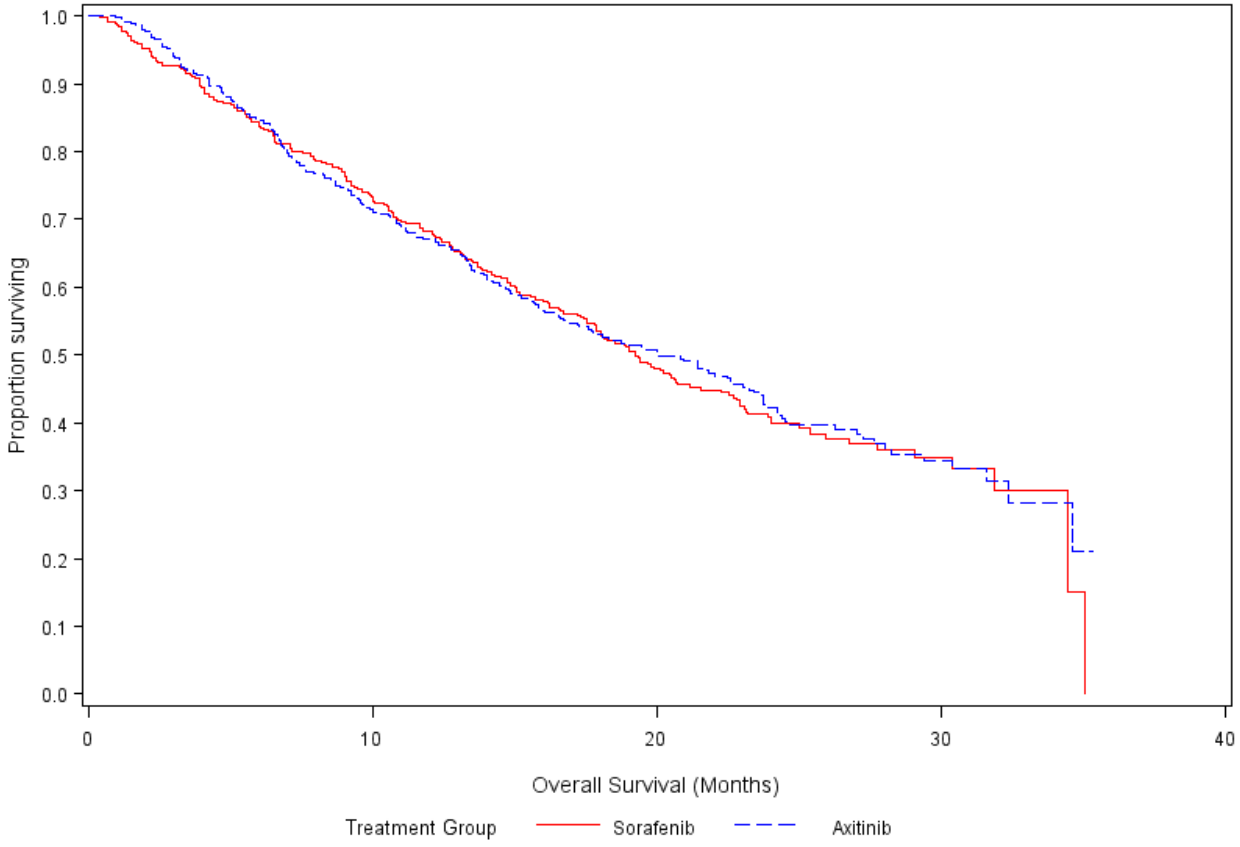


Table 7: Analysis of OS in FAS (Final Analysis, November 1, 2011 Cut-off)

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)	P-value[3]
Axitinib	361	211 (58.45%)	20.1 (16.6, 23.4)	0.969 (0.800, 1.174)	0.3744
Sorafenib	362	214 (59.12%)	19.2 (17.4, 22.3)		

[1]: Kaplan-Meier estimate. [2]: Based on stratified Cox model. [3]: Based on one-sided log-rank test, stratified by ECOG PS and prior therapy, not adjusted for interim analysis.

Figure 5: Kaplan-Meier Plot of OS in FAS (Final Analysis, November 1, 2011 Cut-off)



There were 70 partial responses in the axitinib arm and 34 in the sorafenib arm. There was no complete response in either arm. The objective response rate was 19.4% in the axitinib arm and 9.4% in the sorafenib arm based on the responses assessed by the IRC.

Reviewer's Comments:

1. Lan-DeMets alpha spending function has been used to adjust for multiple OS analyses. According to the Lan-DeMets alpha spending function for O'Brien-Fleming boundary and the actual number of events, the alpha for the interim OS analysis with cut-off date August 31, 2010 and final analysis with the cut-off date November 1, 2011 would be 0.002 and 0.0244, respectively (calculated using East Version 5).
2. There was no adjustment in type I error rate for multiple secondary endpoints. Therefore, p-values for the secondary endpoints are not interpretable.

3.2.8.5. Sensitivity Analyses of PFS

Several sensitivity analyses of PFS are shown below.

1. Table 8 shows the PFS sensitivity analysis based on IRC assessment using the scheduled assessment times.
2. Table 9 shows the PFS sensitivity analysis based on IRC assessment treating the discontinuation due to deteriorating heath status as events.
3. Table 10 shows the PFS sensitivity analysis based on IRC assessment treating censoring due to discontinuation without progression, missed tumor assessments or the start of subsequent anticancer therapy as events.
4. Table 11 shows the PFS sensitivity analysis primarily based on IRC assessment but treating the discontinuation due to investigator-assessed progression with no subsequent scans available as events.
5. Table 12 shows the PFS sensitivity analysis based on IRC assessment in the safety analysis population.
6. Table 13 shows the PFS sensitivity analysis using the earlier of IRC assessed and investigator assessed events or censoring. If IRC and investigator assessed PFS times are same but one of them is an event and the other is censoring, IRC assessment is used.

Table 8: Sensitivity Analysis of PFS Using Scheduled Assessment Time Based on Independent Review in FAS

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)	P-value[3]
Axitinib	361	190 (52.63%)	6.6 (6.5, 8.5)	0.667 (0.545, 0.815)	<0.0001
Sorafenib	362	208 (57.46%)	4.7 (4.6, 6.5)		

[1]: Kaplan-Meier estimate. [2]: Based on stratified Cox model. [3]: Based on one-sided log-rank test, stratified by ECOG PS and prior therapy.

Table 9: Sensitivity Analysis of PFS Treating the Discontinuation Due to Deteriorating Heath Status as Events Based on Independent Review in FAS

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)	P-value[3]
Axitinib	361	196 (54.29%)	6.6 (6.5, 8.3)	0.670 (0.550, 0.817)	<0.0001
Sorafenib	362	214 (59.12%)	4.7 (4.6, 6.5)		

[1]: Kaplan-Meier estimate. [2]: Based on stratified Cox model. [3]: Based on one-sided log-rank test, stratified by ECOG PS and prior therapy.

Table 10: Sensitivity Analysis of PFS Treating Censoring Due to Discontinuation without Progression, Missed Tumor Assessments or the Start of Subsequent Anticancer Therapy as Events

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)	P-value[3]
Axitinib	361	252 (69.81%)	5.6 (4.7, 6.5)	0.655 (0.550, 0.780)	<0.0001
Sorafenib	362	279 (77.07%)	4.0 (3.0, 4.6)		

[1]: Kaplan-Meier estimate. [2]: Based on stratified Cox model. [3]: Based on one-sided log-rank test, stratified by ECOG PS and prior therapy.

Table 11: Sensitivity Analysis of PFS based on IRC Assessment but Treating the Discontinuation due to Investigator-assessed Progression with no Subsequent Scans as Events

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)	P-value[3]
Axitinib	361	221 (61.22%)	6.5 (5.9, 7.8)	0.675 (0.560, 0.814)	<0.0001
Sorafenib	362	241 (66.57%)	4.6 (4.2, 4.9)		

[1]: Kaplan-Meier estimate. [2]: Based on stratified Cox model. [3]: Based on one-sided log-rank test, stratified by ECOG PS and prior therapy.

Table 12: Sensitivity Analysis of PFS based on IRC Assessment in Safety Analysis Set

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)	P-value[3]
Axitinib	359	192 (53.48%)	6.7 (6.3, 8.6)	0.668 (0.547, 0.816)	<0.0001
Sorafenib	355	209 (58.87%)	4.7 (4.6, 5.6)		

[1]: Kaplan-Meier estimate. [2]: Based on stratified Cox model. [3]: Based on one-sided log-rank test, stratified by ECOG PS and prior therapy.

Table 13: Sensitivity Analysis of PFS Using Earlier of IRC-assessed and Investigator-assessed time Scheduled Assessment Time Based on Independent Review in FAS

Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)	P-value[3]
Axitinib	361	197 (54.57%)	6.5 (4.8, 8.3)	0.644 (0.529, 0.784)	<0.0001
Sorafenib	362	218 (60.22%)	4.6 (3.4, 4.8)		

[1]: Kaplan-Meier estimate. [2]: Based on stratified Cox model. [3]: Based on one-sided log-rank test, stratified by ECOG PS and prior therapy.

Reviewer's Comment:

The sensitivity analyses presented above show that the PFS results are robust under different assumptions. The hazard ratios are all very close to that of the primary PFS analysis and all P-values are less than 0.0001.

3.3. Evaluation of Safety

Overall safety profile of axitinib is similar to that of sorafenib. However they had different types of adverse events. For a detailed safety evaluation, please refer to the clinical review of this application.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1. Gender, Race, Age and Geographic Region

Efficacy by gender was analyzed by exploratory analysis of PFS and is presented in Table 14. Efficacy by race was analyzed by exploratory analysis of PFS and is presented in Table 15. More than 97% patients were either White or Asian. Therefore, only those two racial groups are presented in Table 15. Efficacy by age group (<65 years, ≥65 years) was also analyzed by exploratory analysis of PFS and is presented in Table 16. Exploratory analysis of PFS by geographic region (US and Non-US) is presented in Table 17. All PFS analyses in this section are based on independent review. The reported hazard ratios are calculated using unstratified Cox models because subgroup sample sizes are smaller and a stratified analysis may lead to very small number of patients and events per cell.

Table 14: Exploratory Analysis of PFS by Gender

Gender	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)
Female	Axitinib	96	46 (47.92%)	10.1 (6.4, 15.2)	0.429 (0.289, 0.637)
	Sorafenib	104	67 (64.42%)	3.0 (2.8, 4.7)	
Male	Axitinib	265	146 (55.09%)	6.6 (4.8, 8.3)	0.827 (0.656, 1.042)
	Sorafenib	258	143 (55.43%)	4.9 (4.6, 6.5)	

[1]: Kaplan-Meier estimate. [2]: Based on Cox model.

Table 15: Exploratory Analysis of PFS by Race

Race	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)
White	Axitinib	278	157 (56.47%)	6.5 (4.8, 8.3)	0.735 (0.589, 0.919)
	Sorafenib	269	160 (59.48%)	4.7 (4.5, 5.6)	
Asian	Axitinib	77	32 (41.56%)	10.3 (6.4, NE)	0.546 (0.344, 0.867)
	Sorafenib	81	43 (53.09%)	4.7 (2.8, 6.5)	

[1]: Kaplan-Meier estimate. [2]: Based on Cox model. NE: Not estimable.

Table 16: Exploratory Analysis of PFS by Age Group

Age Group	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)
<65 Years	Axitinib	238	133 (55.88%)	6.5 (4.7, 8.3)	0.679 (0.535, 0.861)
	Sorafenib	238	145 (60.92%)	4.7 (3.5, 5.6)	
≥ 65 Years	Axitinib	123	59 (47.97%)	9.3 (6.4, 12.1)	0.696 (0.486, 0.996)
	Sorafenib	124	65 (52.42%)	4.7 (4.3, 6.5)	

[1]: Kaplan-Meier estimate. [2]: Based on Cox model.

Table 17: Exploratory Analysis of PFS by Geographic Region

Geographic Region	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)
US	Axitinib	77	38 (49.35%)	6.7 (4.8, 10.4)	0.616 (0.402, 0.943)
	Sorafenib	92	50 (54.35%)	3.5 (2.8, 6.6)	
Non-US	Axitinib	284	154 (54.23%)	6.8 (6.3, 8.6)	0.707 (0.565, 0.884)
	Sorafenib	270	160 (59.26%)	4.7 (4.6, 6.3)	

[1]: Kaplan-Meier estimate. [2]: Based on Cox model.

Reviewer's Comment:

Axitinib showed improvement over sorafenib across all age groups, gender, race categories and geographic regions with respect to PFS but its improvement in females appears to be much more than in males and its improvement in Asians appears to be more than that in whites.

4.2. Other Special/Subgroup Populations

Exploratory analysis of PFS by ECOG performance status, prior therapy (cytokines or sunitinib) and combination of prior therapy and geographic region are presented in Table 18, Table 19, Table 20, respectively. Since each of temsirolimus and bevacizumab was received as a prior therapy by less than 10% of patients, PFS analyses in those subgroups are not presented here. All PFS analyses are based on independent review. For analysis by combination of prior therapy and geographic region the hazard ratios are obtained using unstratified Cox model. ECOG performance status and prior therapy were two stratification factors at randomization. For analyses by ECOG performance status the hazard ratio was obtained using a Cox model stratified by prior therapy and for analyses by prior therapy the hazard ratio was obtained using a Cox

model stratified by ECOG performance status. The last two analyses mentioned were pre-specified.

Table 18: Exploratory Analysis of PFS by Baseline ECOG Performance Status

ECOG Performance Status	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)
0	Axitinib	197	102 (51.78%)	8.2 (6.4, 11.0)	0.699 (0.532, 0.918)
	Sorafenib	199	110 (55.28%)	5.6 (4.7, 6.4)	
1	Axitinib	164	90 (54.88%)	6.4 (4.5, 8.6)	0.677 (0.508, 0.902)
	Sorafenib	163	100 (61.35%)	4.5 (2.8, 4.7)	

[1]: Kaplan-Meier estimate. [2]: Based on Cox model stratified by prior therapy.

Table 19: Exploratory Analysis of PFS by Prior Therapy

Prior Therapy	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)
Cytokines	Axitinib	126	50 (39.68%)	12.1 (10.1, 13.9)	0.466 (0.321, 0.678)
	Sorafenib	125	69 (55.20%)	6.5 (6.3, 8.3)	
Sunitinib	Axitinib	194	117 (60.31%)	4.8 (4.5, 6.4)	0.744 (0.576, 0.961)
	Sorafenib	195	120 (61.54%)	3.4 (2.8, 4.7)	

[1]: Kaplan-Meier estimate. [2]: Based on Cox model stratified by ECOG performance status.

Table 20: Exploratory Analysis of PFS by Prior Therapy and Region

Prior Therapy and Region	Treatment	Number of Patients	Number (%) Failed	Median in Months[1] (95% CI)	Hazard Ratio[2] Axitinib/ Sorafenib (95% CI)
Cytokines in US	Axitinib	19	6 (31.58%)	15.7 (2.8, NE)	0.713 (0.245, 2.077)
	Sorafenib	19	8 (42.11%)	8.3 (4.7, NE)	
Cytokines outside US	Axitinib	107	44 (41.12%)	12.0 (10.0, 13.8)	0.417 (0.278, 0.623)
	Sorafenib	106	61 (57.55%)	6.5 (4.7, 8.2)	
Sunitinib in US	Axitinib	47	26 (55.32%)	6.3 (3.2, 8.3)	0.576 (0.347, 0.958)
	Sorafenib	64	37 (57.81%)	3.0 (1.6, 4.0)	
Sunitinib outside US	Axitinib	147	91 (61.90%)	4.6 (3.2, 6.4)	0.812 (0.602, 1.095)
	Sorafenib	131	83 (63.36%)	4.6 (2.8, 4.7)	

[1]: Kaplan-Meier estimate. [2]: Based on Cox model. NE: Not estimable.

Reviewer's Comments:

1. All the subgroup analyses presented in this section are considered exploratory or hypothesis generating and no formal inference may be drawn.
2. The PFS improvement in the axitinib arm is consistent across various subgroups.
3. PFS improvement in axitinib over sorafenib appears higher in the subgroup of patients with prior cytokine therapy, especially outside US, than in the subgroup of patients with prior sunitinib therapy.

5. SUMMARY AND CONCLUSIONS

This application is based on one Phase III trial (Study A4061032) and three uncontrolled single-arm Phase II studies (A4061012, A4061035 and A4061023). This review is primarily based on the Phase III study. Study A4061032 was a multicenter, international, randomized, controlled, Phase III study to evaluate the efficacy of axitinib compared to sorafenib in patients with metastatic RCC following failure of one prior systemic first line regimen containing one or more of the following: sunitinib, bevacizumab + IFN-α, temsirolimus, or cytokine(s). Patients were randomized in a 1:1 ratio to receive axitinib at a starting dose of 5 mg twice daily orally with food or sorafenib at a starting dose of 400 mg twice daily orally without food. The randomization was stratified by ECOG performance status (0 vs. 1) and by prior therapy (sunitinib-containing regimens vs. bevacizumab-containing regimens vs. temsirolimus-containing regimens vs. cytokine-containing regimens). The study was initiated on September 15, 2008. The data cut-off date was August 31, 2010.A total of 723 patients were randomized, 361 to axitinib and 362 to sorafenib. Randomized patients were enrolled at 175 centers in 22 countries. There were 169 patients from US. The primary efficacy endpoint was progression-free survival (PFS) as assessed by the independent radiology committee (IRC) review. The secondary efficacy endpoints were investigator-assessed PFS, overall survival (OS), objective response rate (ORR) as assessed by IRC review and duration of response.

The axitinib arm showed statistically significant improvement over sorafenib with respect to PFS as assessed by the IRC in the full analysis set (FAS) [hazard ratio=0.667, 95% confidence interval: (0.546, 0.814), log-rank test stratified by ECOG performance status and prior therapy, one-sided p-value<0.0001]. The study did not show difference in OS between axitinib and sorafenib arms in the FAS [hazard ratio=0.969, 95% confidence interval: (0.800, 1.174), stratified log-rank test, one-sided p-value=0.3744]. However, no adjustment to the level of significance was made for multiple secondary endpoints. Therefore p-values are not interpretable for the secondary endpoints. The objective response rate was 19.4% in the axitinib arm and 9.4% in the sorafenib arm based on the responses assessed by the IRC.

5.1. Statistical Issues and Collective Evidence

1. The initial target enrollment was 540 patients. The applicant increased the sample size to 650 in the Protocol Amendment 4 dated November 16, 2009 and the corresponding Statistical Analysis Plan (version 1.2) dated December 7, 2009. The applicant claims that this change was made due to an underestimation of the dropout rate in the original protocol. However, only 21 patients in the axitinib arm (5.8% of the all axitinib patients) and 37 patients in the sorafenib arm (10.2% of the all sorafenib patients) were censored prior to data cut-off. Changing the sample size alone without any change in the number of events does not affect the error rates.

2. The actual sample size for the study was 723, which is even much higher than 650, the increased sample size that was determined at the time of Protocol Amendment 4. There was no justification provided for such a big difference between the planned and actual sample size. According to the reviewer, it appears that the fast accrual is a possible

reason for accruing many more patients than targeted. The number of PFS events as assessed by IRC was 402 at the final analysis which is very close to the targeted number of events 409. Also, it should be noted that the sample size increase was not because of the guideline for sample size adjustment based on the conditional power as planned in the protocol.

3. The interim analysis for futility was planned to be performed after approximately 204 PFS events (approximately 50% of the total number of events as assessed by the IRC). The actual interim analysis was performed with 289 PFS events (70.7% of the original 409 required events) observed. The applicant claims that because the interim analysis occurring later than planned, the required number of PFS events would need to be increased to 423 in order to maintain 90% power with the same design parameters. The applicant did not amend the protocol or the Statistical Analysis Plan to include this modification. Applicant's claim about the need to increase the number of events was not correct. Based on the reviewer's calculation using East version 5, the number of events should not be more than 414. The number of PFS events at the final analysis was 428 based on investigator's assessment and 402 based on IRC.

4. Type I error rate has not been adjusted for analysis of multiple secondary endpoints. Therefore, p-values for the secondary endpoints are not interpretable.

5. Patient-reported outcome endpoints are subject to bias in an open-label study. Therefore, this review considers those endpoints in this open-label trial exploratory.

6. Although the subgroup analyses by the stratification factors (ECOG performance status and prior therapy) were pre-specified, there was no adjustment was made in the Type I error rate for multiple subgroup analyses. Therefore, these subgroup analyses are considered exploratory by the reviewer.

7. The PFS improvement in the axitinib arm is consistent across various subgroups.

8. There were some problems with submission datasets most of which were corrected by the applicant in subsequent submissions.

5.2. Conclusions and Recommendations

The applicant has submitted results from one multicenter, phase III, randomized, open-label, clinical trial (Study A4061032) comparing axitinib, a new molecular entity (NME), to sorafenib in patients with metastatic RCC following failure of one prior systemic first line regimen containing one or more of the following: sunitinib, bevacizumab + IFN-α, temsirolimus, or cytokine(s). The axitinib arm showed statistically significant improvement over sorafenib in progression-free survival (PFS) as assessed by independent radiology committee in all randomized patients. However, the axitinib arm did not show statistically significant improvement with respect to overall survival (OS). The application was discussed at the Oncologic Drug Advisory Committee meeting on December 7, 2011. The committee voted

unanimously in favor of axitinib to the question whether the benefit-risk ratio of axitinib is favorable. The statistical results provide adequate evidence to support the PFS claim proposed in the NDA.

SIGNATURES/DISTRIBUTION LIST

Primary Statistical Reviewer: Somesh Chattopadhyay, Ph.D.
Date: January 9, 2012

Concurring Reviewer(s): Shenghui Tang, Ph.D., Team Leader
Rajeshwari Sridhara, Ph.D., Director

Statistical Team Leader: Shenghui Tang, Ph.D.

Biometrics Division Director: Rajeshwari Sridhara, Ph.D.

cc:
HFD-150/Ms. Lisa Skarupa
HFD-150/Dr. Amy McKee
HFD-150/Dr. John Johnson
HFD-711/Dr. Somesh Chattopadhyay
HFD-711/Dr. Shenghui Tang
HFD-711/Dr. Rajeshwari Sridhara
HFD-700/Ms. Lillian Patrician

CHECK LIST

Number of Pivotal Studies: 1

Trial Specification

Protocol Number (s): A4061032
Protocol Title (optional): Axitinib (AG-013736) as second line therapy for metastatic renal cell cancer: Axis trial
Phase: 3
Control: Active Control
Blinding: Open-Label
Number of Centers: 175
Region(s) (Country): Australia (5 centers), Austria (2 centers), Brazil (4 centers), Canada (6 centers), China (7 centers), France (10 centers), Germany (7 centers), Greece (2 centers), India (5 centers), Ireland (1 center), Italy (13 centers), Japan (18 centers), Republic of Korea (6 centers), Poland (4 centers), Russian Federation (7 centers), Singapore (1 center), Slovakia (3 centers), Spain (9 centers), Sweden (3 centers), Taiwan (4 centers), United Kingdom (UK, 10 centers), and United States (US, 48 centers); an additional 19 centers were shipped study drug (including 4 centers in The Netherlands), but did not enroll any patients.
Duration: Until disease progression, intolerable adverse drug reactions, or withdrawal of consent
Treatment Arms: Experimental: axitinib, Control: sorafenib
Treatment Schedule: Axitinib: 5 mg administered orally twice daily, Sorafenib: 400 mg administered orally twice daily without food
Randomization: Yes
Ratio: 1:1
Method of Randomization: Central via an IVRS with stratification
Stratification Factors: ECOG performance status (0 vs. 1) and prior therapy (i.e., sunitinib-containing regimens vs. bevacizumab-containing regimens vs. temsirolimus-containing regimens vs. cytokine-containing regimens)
Primary Endpoint: Progression-free survival (PFS)
Primary Analysis Population: ITT
Statistical Design: Superiority
Adaptive Design: Yes (sample size adjustment based on interim analysis results; adjustment was not needed)
Primary Statistical Methodology: Staratified log-rank test
Interim Analysis: Yes
If yes:
No. of Times: 1 for PFS futility and 2 for OS efficacy
Method: O'Brien-Fleming
α Adjustment: Yes
α Spending Function: Lan-DeMets alpha spending function for O'Brien-Fleming boundary

DSMB: Yes

Sample Size: 723 (planned 650)

Sample Size Determination: Was it calculated based on the primary endpoint variable and the analysis being used for the primary variable? Yes

Statistic = Log-rank

Power= 90%

Δ= Hazard ratio of (median PFS of 7 months in experimental arm vs. 5 months in control arm

α = one-sided 0.025

- Was there an **Alternative Analysis** in case of violation of assumption; e.g., Lack of normality, Proportional Hazards Assumption violation. NA
- Were there any major changes, such as changing the statistical analysis methodology or changing the primary endpoint variable? No
- Were the **Covariates** pre-specified in the protocol? Yes.
- Did the Applicant perform **Sensitivity Analyses**? Yes.
- How were the **Missing Data** handled? For PFS, the applicant followed the Oncology Endpoint Guidance. For OS, censoring was used.
- Was there a **Multiplicity** involved? Yes.

  If yes,

  Multiple Arms (Yes/No)? No.

  Multiple Endpoints (Yes/No)? Yes.

  Which method was used to control for type I error? There was no adjustment for multiple secondary endpoints.
- **Multiple Secondary Endpoints**: Are they being included in the label? If yes, method to control for type 1 error. No.

Were Subgroup Analyses Performed (Yes/No)? Yes.

- Were there any **Discrepancies** between the protocol/statistical analysis plan vs. the study report? Yes. The applicant intended to increase the number of PFS events without amending the protocol.
- Overall, was the study positive (Yes/No)? Yes.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

SOMESH CHATTOPADHYAY
01/11/2012

SHENGHUI TANG
01/11/2012

RAJESHWARI SRIDHARA
01/11/2012

STATISTICAL REVIEW AND EVALUATION

Biometrics Division: VI (HFD-705)

NDA NO.:	202-324
SERIAL NO.:	S-000
DATE RECEIVED BY THE CENTER:	April 14, 2011
DRUG NAME:	Axitinib
DOSAGE FORM:	Tablets
INDICATION:	Oncology
SPONSOR:	Pfizer, Inc.
DOCUMENTS REVIEWED:	Submission dated April 14, 2011 and subsequent responses
NAME OF STATISTICAL REVIEWER:	Meiyu Shen, Ph.D. (HFD-705)
PROJECT MANAGER:	Don Henry

_____

Meiyu Shen, Ph.D., Mathematical Statistician

Concur: _____

Yi Tsong, Ph.D.
Deputy Director, DBVI

Distribution: NDA 202-324
HFD-705/Y. Tsong, Ph.D.
ONDQA/Don Henry

TABLE OF CONTENTS

EXECUTIVE SUMMARY OF STATISTICAL FINDINGS

1.1 Purpose of this statistical consultation

Pfizer submitted manufacturing process development and analytical procedures development and validation in NDA 202-324. In this submission, the sponsor requested the regulatory flexibility of the method operable design region (MODR) for high performance liquid chromatography in Validation of Analytical Procedure Drug substance (3.2.S.4.3) and drug product in Validation of Analytical Procedures (3.2.P.5.3). The regulatory flexibility means that changes within MODR or DS are non-regulatory change.

On May 17, 2011, Division of Biometrics VI received the official request for the following consult: "perform statistical evaluations of the proposed design space for the manufacturing process." In this review document, the statistical reviewer reviewed the original submission and the sponsor's responses to September 26-2011 FDA Query.

(b) (4)

8 Page(s) has been Withheld in Full as B4 (CCI/ TS) immediately following this page

---

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

MEIYU SHEN
12/11/2011

YI TSONG
12/12/2011

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 202324 **Applicant: Pfizer Inc.** **Stamp Date: April 14, 2011**

Drug Name: Axitinib (Inlyta) **NDA/BLA Type: Original NDA**

On **initial** overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	X			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)		X		The original protocol for the pivotal study has not been submitted. Only the final version is submitted.
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).	X			
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).	X			Define.pdf file is unclear.

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? ___Yes_____

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	X			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	X			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.	X			
Appropriate references for novel statistical methodology (if present) are included.			X	
Safety data organized to permit analyses across clinical trials	X			

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

in the NDA/BLA.				
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.			X	

Somesh Chattopadhyay — May 13, 2011

Reviewing Statistician — Date

Shenghui Tang — May 13, 2011

Supervisor/Team Leader — Date

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

SOMESH CHATTOPADHYAY
05/25/2011

SHENGHUI TANG
05/25/2011