TABLE OF CONTENTS

STATISTICAL ANALYSIS PLAN ADDENDUM

(Version 1.0)

Protocol Title:	A Phase III Double-blinded, Placebo Controlled Study of Xilonix™ for Improving Survival in Metastatic Colorectal Cancer
Sponsor:	XBiotech USA Inc
Protocol Number:	2012-PT023
Protocol Version/Date:	Version 3.7 / June 5th, 2017
IND Number:	114759

APPROVALS

XBiotech USA, Inc.

Mark J. Jaros, Ph.D. — 18 AUG 2017 — Date
Statistical Consultant

Anna Wang — 21 Aug 2017 — Date
Clinical Data Manager

Michael Stecher, M.D. — 18 Aug 2017 — Date
Medical Director

Document History

Version	Reason(s) for change	Date
V1.0	Original document	August 18th 2017

This statistical analysis plan (SAP) addendum describes hard-coding decisions for protocol 2012-PT023, A Phase III Double-blinded, Placebo Controlled Study of Xilonix™ for Improving Survival in Metastatic Colorectal Cancer. This addendum was finalized after the SAP was finalized.

The date of death for the following patients need to be hard-coded in SAS program to derive the overall survival endpoint. The detailed rationales for these decisions are described below, and these decisions were made while the database was still blinded.

In general, the issues were isolated into three categories: Partial Dates, Missing Date of Death, and Randomization in Error.

- For partial dates, ie those that contain month and year, a conservative approach was taken and the 1st of the month was imputed in these cases.
- For missing date of death, the last date of contact with the patient was used as the date of death.
- The EDC assigns a randomization number whenever the randomization option is selected on that page. There were few of these data entry errors (N=6) in the study, and in each case the site confirmed the error. These patients are considered to be screen failures.

The following cases have confirmed death, but the date of death is missing (unknown):

- Patient 3301233005, this patient died based on the survival data sheet, but the date of death is missing. In study discontinuation data sheet, this patient had an adverse event on 08Apr2016, which was used as the date of death in the interim analysis dataset. In the survival data sheet, the last available assessment date is 13Jun2016. As the patient was known to be alive at this time, the last available contact date should be used.
- Patient 3903233019, this patient died based on the survival data sheet, but the date of death is missing. In study discontinuation data sheet, this patient had date for "Other" on 05-09-2016, which was used as the date of death in the interim analysis dataset. In the survival data sheet, the last available assessment date is 29Jun2016 for ECOG assessment. As the patient was known to be alive at this time, the last available contact date should be used.
- Patient 3802233004, this patient died based on the survival data sheet, but the date of death is missing. In study discontinuation data sheet, this patient had date for "Removed from Study by PI" on 22-Nov-2016. In the survival data sheet, the last available assessment date is 21Mar2017 for ECOG assessment. As the patient was known to be alive at this time, the last available contact date should be used.

The following cases do not need hard-coding, but imputations will be done for partial date of death:

- Patient 3905233015, the death date has missing day part, Mar-2017, this date will be imputed as 01-Mar-2017.
- Patient 3909233001, the death date has missing day part, Feb-2017, this date will be imputed as 01-Feb-2017.
- Patient 3905233011, the death date has missing day part, Jun-2016. In the interim analysis dataset, the death date was set as 06-Jun-2016. After further clarification, this is the date that the site completed the discontinuation page in the EDC. In a query response, the site noted that they did not know the exact date of death, but they knew the patient died in June of 2016. As the June 6th date is not the last available contact date with the patient, rather the date the site entered discontinuation information, this date should not be used. This case is being classified as a partial date, therefore the date of 01-Jun-2016 will be imputed.
- Patient 3909233003, the death date has missing day part, Dec-2016. In the interim analysis dataset, the date of death was set as 01-DEC-2016. This is consistent with the decision to use the 1st of the month and therefore 01-DEC -2016 will be used as the date of death.

The following patients were screen failures, however the site erroneously clicked on the randomization button in the EDC. Once this happens, a randomization number is assigned and this can't be undone. Because these patients did not meet eligibility criteria and the site confirmed the error, these patients are not included in the data analysis.

- 1027233009
- 3909233006
- 1025233003
- 1029233008
- 1069233002
- 3301233002

STATISTICAL ANALYSIS PLAN

(Version 1.1)

Protocol Title:	A Phase III Double-blinded, Placebo Controlled Study of Xilonix™ for Improving Survival in Metastatic Colorectal Cancer
Sponsor:	XBiotech USA Inc
Protocol Number:	2012-PT023
Protocol Version/Date:	Version 3.6 / November 15th, 2016
IND Number:	114759

APPROVALS	
XBiotech USA, Inc.	
Prasant Mohanty, **Director of Biostatistics**	18 NOVEMBER 2016 **Date**
Dawn McCollough **VP, Clinical Operations**	18 November 2016 **Date**
Anna Wang, **Clinical Data Manager**	18 NOV 2016 **Date**
Michael Stecher, M.D. **Medical Director**	18 NOV 2016 **Date**

Document History

Version	Reason(s) for change	Date
V1.1	Primary analysis population changed from mITT (patients receiving at least one dose of study drug) to ITT (all randomized patients)	November 18, 2016

Preface

This statistical analysis plan (SAP) describes the planned analysis and reporting for protocol 2012-PT023, A Phase III Double-blinded, Placebo Controlled Study of Xilonix™ for Improving Survival in Metastatic Colorectal Cancer.

The contents of this SAP is intended to be sufficiently detailed to meet the requirements of International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH).

This SAP is prepared by the sponsor's statistician who is blinded to the treatment allocation and will remain so until database is locked and final data are extracted for analysis.

For details on the conduct of this study and to learn about the operational aspects of the clinical assessments, please read the trial protocol (referenced above).

Abbreviations

Abbreviation	Definition
ANOVA	Analysis of Variance
ANCOVA	Analysis of Covariance
BSC	Best Supportive Care
BOR	Best Overall Response
CHP	Cox Proportional Hazards mode
CI	Confidence Interval
CR	Complete Response
DCR	Disease Control Rate
DEXA	Dual-energy X-ray absorptiometry
ECOG	Eastern Cooperative Oncology Group
EORTC QLQ-C30	European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30
HR	Hazard Ratio
ICH	International Conference on Harmonisation
IDMC	Independent Data Monitoring Committee
IQR	Inter Quartile Range
irRC	Immune-related Response Criteria
ITT	Intent to Treat
IWRS	Interactive Web Response System
LBM	Lean Body Mass
MedDRA	Medical Dictionary for Regulatory Affairs
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
OS	Overall survival
ORR	Objective Response Rate
PFS	Progression Free Survival
PR	Partial Response
PT	Preferred Term
RR	Response Rate
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Stable Disease
SD	Standard Deviation
SOC	System Organ Class

Table of Contents

1 Introduction

1.1 Background:

In the setting of refractory, metastatic disease a complete resolution of tumor burden is not a reasonable expectation. Instead, the primary goal of anti-tumor therapy at this stage is to eliminate or reduce the symptomatic effects of the tumor, while trying to prolong survival for as long as possible. Due to treatment related morbidity however, few treatment modalities are ideal for this objective. Even with the most recent targeted agents (such as multi-kinase inhibitors), drug related toxicities frequently lead to relatively short treatment durations. With discontinuation of therapy, disease progression is uncontrolled and prognosis is poor.

New agents that control disease progression—while improving tumor-related symptoms, rather than causing significant therapy related morbidity—are vitally needed to treat patients with advanced cancer, including those with colorectal cancer. An approach has been taken to develop such an agent using a monoclonal antibody to block the chronic inflammation involved in both malignant disease progression and constitutional symptoms.

Xilonix™ is expected to inhibit tumor growth and metastasis by interrupting crucial signals that drive angiogenesis and invasiveness. The antibody therapy may also block tumor microenvironment infiltration by leukocytes (such as myeloid suppressor cells) that suppress antitumor immunity, enabling better host immune control of the disease. In addition to local effects on the tumor, Xilonix™ is expected to work systemically to correct the metabolic dysregulation, fatigue and anxiety mediated by chronic inflammatory signaling to the central nervous system. The use of this antibody monotherapy to target chronic inflammation is proposed as a safe, effective treatment for patients with metastatic colorectal cancer.

2 Study Design

This is a phase III, multicenter, double blind, randomized, placebo controlled pivotal trial of the True Human monoclonal antibody MABp1 in subjects with metastatic colorectal cancer who are refractory to standard therapy.

- Enrolled subjects will be randomized (2:1) to receive either MABp1 plus best supportive care (BSC) or placebo plus BSC.
- BSC is defined as those measures intended to provide palliation of symptoms and improve quality of life. This includes, but is not limited to, antibiotics, anti-emetics, narcotics, and parenteral nutrition.
- Subjects randomized to MABp1 or placebo will receive the study drug via intravenous infusion once every 2 weeks (one cycle).

3 Study Endpoints:

3.1 Primary Efficacy Endpoint:

- Overall survival (OS) will be the primary endpoint of this study, survival time will be defined as the duration from the date of randomization until death. Subjects who are alive at the data inclusion cutoff date for the analysis will be censored at that date and survival time will be defined as time from randomization to censor date. Patients lost to follow-up and untraceable for survival status will be censored on the last date subject was known to be alive or lost to follow-up. The overall survival between the Xilonix™ + BSC group and placebo + BSC group will be compared using an unadjusted log-rank test.

3.2 Secondary Efficacy Endpoints:

Secondary efficacy variables will include 1) change in lean body mass (LBM) measured by dual-energy X-ray absorptiometry (DEXA) scans, 2) change in function, symptoms and overall quality of life assessed through the cancer-specific EORTC QLQ-C30[1] questionnaire, 3) change of platelet counts, 4) progression free survival (PFS), 5) objective response rate (ORR) 6) disease control rate (DCR)
The change from baseline on LBM, EORTC score, and platelet count will be calculated by subtracting the baseline (screening) value from that at follow-up (cycle 5 day 1 or off-study visit for patients who completed at least cycle 3 day 1).

Quality of life (QoL) in patients participating in the trial will be evaluated using the European Organization for Research and Treatment of Cancer (EORTC) version 3.0 questionnaire (Appendix). The EORTC QLQ-C30 questionnaire incorporates nine multi-item scales: five functional scales (Physical, Role, Cognitive, Emotional and Social); three symptom scales (Fatigue, Pain and Appetite loss); and a Global Health Status/QoL scale. Each item, excepting Global Health Status, is answered on a four-point scale: "not at all", "a little", "quite a bit", and "very much". Response to Global Health Status is measured on a 1 to 7 scale, 1 being "very poor" and 7 being "excellent". Using EORTC specified methodology (Appendix); responses to all the scales will be linearly transformed to reach the scale range 0 to 100.

Immune Related Response Criteria: Tumor response will be assessed by using immune-related response criteria (irRC). The imaging service provider (core laboratory) will acquire images from clinical sites. Review of image data will be conducted in the core lab by qualified and trained radiologists blinded to treatment status. The procedures used for CT/MRI image analysis, reporting of tumor response data and storage and archiving of images will follow the standards laid out in the prospectively determined Imaging Review Charter.

irRC Evaluation:
The detailed evaluation methodology is documented in the Imaging Review Charter. Briefly, for each imaging session, the sum of the products of the two largest perpendicular diameters (SPD) of all index lesions (5 lesions per organ, up to 10 visceral and 5 cutaneous lesions with

a minimum size of ≥5 x 5 mm) will be calculated. Index lesions will be identified at baseline, and will be followed throughout the patient's participation in this study.
At each subsequent imaging assessment, new lesions (defined as a lesion with > 5 x 5 mm measurements) will have their products of diameters measured for use in calculating tumor burden. Tumor burden will be calculated by summing the SPD of target lesions with the SPD of the new measurable lesions.

Table 1: irRC Response Definitions:

Complete Response (irCR):	Disappearance of all lesions in two consecutive observations not less than 4 weeks apart
Partial Response (irPR):	≥50% decrease in tumor burden compared with baseline in two observations at least 4 weeks apart
Stable Disease (irSD):	50% decrease in tumor burden compared with baseline cannot be established nor 25% increase compared with nadir
Progressive Disease (irPD):	At least 25% increase in tumor burden compared with nadir (at any single time point) in two consecutive observations at least 4 weeks apart
Not Evaluable (NE)	Progression has not been documented and one or more index lesions have not been assessed or have been assessed using a different method than baseline that makes comparability impossible.

Table 2: Derivation of irRC Overall Response

Overall lesions response at each assessment			
Index and new lesions (tumor burden), * %	Non-Target Lesions	New Lesions	Overall Response
↓100	Absent	Absent	irCR †
↓100	Stable	Any	irPR†
↓100	Unequivocal progression	Any	irPR†
↓> 50	Absent/Stable	Any	irPR†
↓> 50	Unequivocal progression	Any	irPR†
↓< 50 to < 25↑	Absent/Stable	Any	irSD
↓< 50 to < 25↑	Unequivocal progression	Any	irSD
> 25	Any	Any	irPD†

* Decreases assessed relative to baseline, including measurable lesions only (≥5 × 5 mm).
†Assuming response (irCR and irPR) and progression (irPD) are confirmed by a second, consecutive assessment at least 4 weeks apart.

Best overall response (BOR), determined from the sequence of cycle responses assessed, will be used for analyzing ORR and DCR.

Progression Free Survival (PFS): This will be defined as time from randomization to progression of disease or death as defined in Table 1 below, whichever is earlier. Patients surviving without objective disease progression at the end of the data inclusion cutoff date will be censored. If there is no baseline disease progression assessment, then PFS will be censored at randomization. An adequate tumor assessment is defined as a set of radiologic images, which has been transmitted to the central imaging vendor for the purposes of assessing change in tumor measurements according to the immune related response criteria. In order to be considered adequate, the images should include assessment of chest, abdomen and pelvis using the same imaging modality as the baseline scan and should include all target and non-target lesions.

Table 3: Definitions for Assessment of PFS

Situation	**Date of Progression or Censoring**	**Outcome**
No baseline assessment	Randomization	Censored
Progression documented between scheduled visits	Next scheduled visit	Progressed
No progression	Date of last visit with adequate assessment	Censored
Investigator claim of clinical progression	Scheduled visit (or next scheduled visit if between visits)	Progressed
Treatment discontinuation for toxicity or other reason	Date of last visit with adequate assessment	Censored
New anticancer treatment started with no claim of progression	Date of last visit with adequate assessment	Censored
Death before first PD assessment	Date of death	Progressed
Death between adequate assessment visits or after patient misses one assessment visit	Date of death	Progressed
Death after an extended lost-to-follow-up time (two or more missed assessments)	Last visit with adequate assessment	Censored

Objective Response Rate (ORR): The ORR will be estimated by dividing the total number of patients with confirmed complete response (CR) and partial response (PR) by the total number of patients in the ITT population.

Disease Control Rate (DCR): The DCR will be estimated by dividing the total number of patients with confirmed CRs, PRs and stable disease (SD) by the total number of patients in the ITT population.

3.3 Exploratory Endpoint:

Clinical Response Rate (CRR) will be the exploratory endpoint for this study. This will be a composite measure assessing the change in LBM, fatigue, pain, and appetite from baseline to week 8. Clinical response will be defined as, 1) Improvement or stabilization (≥ 0 kg change) of LBM as assessed by DEXA scan, and 2) Improvement or no worsening (≤ 0 score point change; low scores reflect better functioning and lower symptom distress) on any two of the three symptom scale measures (fatigue, pain, appetite) of EORTC QLQ-C30. Patients missing the follow-up assessment will be considered non-responders.

3.4 Safety Endpoints:

Safety endpoints will be evaluated by monitoring adverse events from clinical and laboratory reporting. Adverse events will be graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE, version 4.0). Significant trends in the distribution and severity of the adverse events across the study groups will be assessed.

4 Independent Data Monitoring Committee (IDMC):

An Independent Data Monitoring Committee (IDMC) will be formed to safeguard the interests of trial participants, assess the safety and efficacy of the interventions during the trial, and monitor the overall conduct of the clinical trial. The committee will meet after every 200 subjects enrolled (or 6 months, whichever occurs first) until the first interim analysis to monitor early evidence of treatment harm, and make any appropriate recommendations. After the first interim analysis, the meeting will occur at each subsequent analysis or every year, whichever occurs first. IDMC recommendations will include whether to continue, suspend, or otherwise modify enrolment. Monitoring for early toxicity will be continuous until the first interim analysis, after which this monitoring will occur at each subsequent scheduled interim analysis or every year, whichever occurs first. Full details of the role and responsibilities of the IDMC are described in the IDMC charter.

5 Sample Size and Power

Prior research has reported a median survival of 4.6 months among refractory, metastatic colorectal cancer patients with BSC in a comparable patient cohort. On the other hand, as observed in the colorectal cancer cohort of our phase I/II trial, the median survival was 6.7 months for all enrolled colorectal cancer patients and 8.7 months for the per-protocol population. Using a conservative approach, the current study is designed to detect at least 30% improvement in median OS in treatment arm (6.0 months, hazard rate in treatment arm (h1)=0.116)) compared to the control arm (4.6 month, hazard rate in control arm (h0)=0.151)). We propose to test the null hypothesis for hazard ratio (h1/h0) =1 and alternative hazard ratio ≤ 0.767. A total of 552 deaths will be required to detect the 30% increase in median survival (from 4.6 to 6.0 months, hazard ratio 0.767) with one-sided Type I error probability (alpha) of 0.025, and 80% power (1-

beta=0.80). With a 2:1 allocation ratio and 5% oversampling, a total of 600 subjects (400 and 200 in treatment and control arm respectively) with 18 month follow-up of the overall study population will provide the required number of events (deaths). The survival curves will be estimated with the Kaplan–Meier method and compared for statistical significance by an unadjusted log-rank test.

5.1 Interim Efficacy Analyses:

Two interim efficacy analyses will be conducted before the final analysis of overall survival. As determined based on alpha spending function of O'Brien and Fleming sequential group design (O'Brien PC, Fleming TR, 1979), survival analyses will be performed after 276 (50%), 414 (75%) and 552 (100%) deaths at the respective stages. After the required number of events for the specific interim analysis has occurred, the database will be frozen for planned statistical analysis.

The criteria for the early termination for efficacy (rejection of null hypothesis) or to accept the null hypothesis will be based on the group sequential design. If the test statistics crosses the pre-specified boundaries for type I error (alpha cut-off 0.0029, 0.0121, and 0.025 at the first, second interim or final analysis respectively) the trial will stop for efficacy. Otherwise the trial continues to the next stage.

The cumulative p value for the acceptance boundary (beta) will be 0.25 and 0.08 at the first and second interim respectively. If efficacy is established at an interim analysis, enrollment will be stopped and the control group will be allowed to crossover. All subjects will be followed up until death, loss to follow-up, or termination of the study.
SAS codes for O'Brien and Fleming sequential group design will be included in the Interim Analysis Plan. TFL shells for interim analysis are provided in the Appendix section. The decision rule with interim monitoring will be as follows:

1. If type I error probability (alpha) $\leq$the above specified cumulative alpha level at the given stage, trial will stop for efficacy
2. Else, if the p value $\geq$the above specified beta levels the trial will stop for accepting null hypothesis
3. Else the trial will continue to the next stage
4. At final analysis, if alpha is $\leq$0.025 then declare efficacy; otherwise accept the null hypothesis.

The Trial Statistician, who is independent from the study statistician, will conduct these interim analyses as well as the end-of-study final analysis. The Trial Statistician will be responsible for presenting the unblinded interim data to the IDMC. Prior to completion of the trial, unblinded access to the interim efficacy data will only be available to the IDMC for review and the committee has the responsibility of making recommendations to the sponsor based on this data.

6 Randomization Strategy

A central randomization scheme with Interactive Web Response System (IWRS) will be employed to facilitate effective randomization and allocation concealment. The scheme will involve a block randomization technique, randomly assigning participants within blocks based on a 2:1 allocation ratio to the Xilonix™ and placebo arm respectively. After confirmation of eligibility, the study coordinators at the clinical site will complete the randomization eCRF indicating the patient has met inclusion criteria and is ready to be randomized. This step will initiate an IWRS procedure, which will assign the next available sequential randomization number specifying the arm allocation. The randomization sequence will be generated using Oracle Clinical Remote Data Capture (OC RDC) application (Oracle Corporation, Redwood City, CA, USA). The randomization code and the study arm assignment will be maintained inside the RDC system. The randomization will not be stratified on any factor.

7 Statistical Methods

7.1 Analysis Populations:

The primary efficacy analysis will be conducted on an intent-to-treat (ITT) basis. The ITT population will consist of all randomized patients. Patients will be grouped according to treatment group.

Some major changes were made to protocols version V3.0 onwards. Prior to V3.0 this trial was an open label, 1:1 random allocation study where patients received MABp1 (3.75 mg/kg) plus BSC or megestrol acetate plus BSC. Some of the major changes to protocols V3.0 onwards were a double blind placebo controlled design with 2:1 allocation ratio. Dosing regimen also changed from 3.75 mg/kg to 7.5 mg/kg of MABp1. Given the nature of change to the study design patients enrolled on protocols prior to V3.0 modification will not be included in the analysis.

A secondary analysis of the efficacy endpoints will be performed on Per Protocol (PP) patients. The PP population will be defined as those patients in the ITT set who are compliant with the study protocol. Patients will be grouped according to treatment groups.

The Safety analysis will be based on the Safety Population, defined as enrolled patients receiving at least one dose of study drug. Patients will be grouped according to actual treatment received.

7.2 Incomplete Follow-up:

Patients discontinued or lost prior to cycle 3 visit (and therefore not having the second DEXA), or those discontinuing between cycles 3 and 5 but did not complete off-study visit, will not be included for secondary efficacy outcome assessment for absolute change in LBM, or EORTC endpoints.

For the PFS assessment, patients discontinued or lost without evidence of tumor progression will be censored according to Table 1 .

7.3 Missing Data:

Every effort will be made to identify and recover the missing data from the centers. If the data remained missing following such efforts, and contributes 2% or more of total number of patients for the continuous co-variates selected for the Cox model, it will be imputed through multiple imputations (using SAS PROC MI, pooled by PROC MIANALYZE option). Sensitivity analysis will be performed to assess the effect of imputation.

The primary efficacy analysis will be performed using the available data, no imputation will be made. Also the missing data for primary or secondary outcome variables, except for platelet counts, will not be imputed. For platelet counts missing on cycle 5 visit, the available value on cycle 3 or cycle 4 visit will be carried forward. Imputation of missing data will not be carried out for descriptive analyses.

7.4 Outliers

No formal outlier tests are planned. Normally distributed values that are outside 3 standard deviations from the mean will be queried and verified with source prior to database lock. Tukey's 3-inter-quartile range (IQR) principle[2] for outlier detection will be utilized for identifying extreme values in platelet count and laboratory measures. Values outside 3 times of 5^{th} and 95^{th} quantile range will be considered as extreme values. If extreme values are present, a sensitivity analysis will performed after removing the extreme values.

7.5 Derived and Computed Variables:

Patient's age (in years) at randomization is a derived field in the remote capture database, which is rounded to an integer value after subtracting date of birth from the randomization date and dividing the results by 365.25 days.

Change from Baseline
For the continuous variables or variables with linear transformation (e.g. EORTC score), the change at follow-up will be calculated by subtracting the value on screening visit from that on the follow-up visit (week 8 or end of study visit, as appropriate). Percent change will be computed as ([post-value – pre-value]/pre-value)*100.

8 Statistical Analyses

8.1 Enrollment and Disposition

Subject enrollment and disposition will be illustrated on a flow diagram (following CONSORT format) showing the randomization assignment, any possible loss to follow-up, and the ITT, PP and Safety population. Data will be described using incidence rates (number of patients, number

of events, and percent). A table showing enrollment by site and country will be provided. This table will include the number of patients randomized, the number completing each dosing cycle, and major reason for discontinuation, grouped by treatment arm.

A listing of all patients discontinued from the study after randomization with specific reason for discontinuation, cumulative dose (for the Xilonix™ group), and duration on study before discontinuation, broken down by enrolling site and treatment group will be provided.

8.2 Patient Characteristics and Baseline Comparisons

The ITT, PP and safety populations will be used for analysis of demographic and baseline findings. Data will be summarized for each treatment group, i.e. Xilonix™ and placebo arm. Descriptive analysis will be performed summarizing age, gender, Eastern Cooperative Oncology Group (ECOG) status, co-morbidities, body mass index (BMI), cancer therapy, and other relevant baseline risk factors by treatment group. Baseline biochemistry, hematology data will be summarized in separate tables.

The number and percentage of patients reporting prior therapies will be provided by type of therapy (surgery or radiation).

Medication start and stop date will be compared to the date of first infusion of study drug, and will be classified as Prior, Prior and Concomitant, or Concomitant. The number and percentage of patients using prior, prior and concomitant, and concomitant antineoplastic medications will be presented by treatment group. WHO Drug Dictionary, September 2016 will be used for coding of concomitant medications

Continuous variables will be reported as mean ± standard deviation (SD), median, inter-quartile range (IQR), and minimum and maximum values. The categorical variables will be reported as number of cases (n) and percentage. The baseline characteristics will be compared across the treatment groups for overall difference by using independent-samples t-test for continuous variables and Fisher's exact for categorical data. Normality of the analysis variables will be tested using Shapiro-Wilks test. If the Shapiro-Wilks test is significant, continuous data will be compared using Wilcoxon (for paired data) and Mann-Whitney nonparametric test. A two-tailed p value of <0.05 will be considered statistically significant for the descriptive summary statistics.

8.3 Protocol Compliance and Follow-up

A summary of scheduled visit compliance will be provided by treatment groups. The frequency distribution of drop outs and early terminations will be provided.

Protocol deviations will be grouped into major categories based on relation to exclusion, inclusion criteria, and conduct of trial, and will be summarized by enrolling centers. The categorization will be completed prior to unblinding.

8.4 Efficacy Evaluation

The primary efficacy analysis will be conducted for the ITT and PP population, as defined in section 6.1. Patients will be grouped according to randomized treatment.

8.4.1 Primary Efficacy Endpoint

Subjects who are alive at the data inclusion cutoff date for the analysis will be censored at that date and survival time will be defined as time from randomization to censor date. Patients lost to follow-up and untraceable for survival status will be censored on the last date subject was known to be alive or lost to follow-up. A censoring variable will be created identifying a censored observation. The primary endpoint, overall survival, will be summarized by Kaplan-Meier method and compared between the treatment groups using an un-adjusted log-rank test. The median survival, and cumulative survival probability along with 95% confidence interval (CI) calculate by log-log transformation and 25^{th} and 75^{th} percentiles will be reported according to randomization arms. Kaplan-Meier curves will be generated plotting the estimated survival density function over time. Number of patients at risk on various time intervals will be presented on the curve. The type I error rate to be used for determining statistical significance of the primary outcome is defined in section 5.1 "Interim Efficacy Analyses".

The primary efficacy analysis will also be conducted for the PP population.

A secondary supportive analysis will be performed comparing survival probabilities across treatment groups, stratified separately by age (dichotomizing at median), gender, ECOG performance status, KRAS mutation status and prior treatment with regorafenib, and number of prior antineoplastic therapies. Survival probabilities and 95% CI for the 25^{th} , median and 75^{th} percentiles will be reported by treatment groups and stratum. No multiplicity adjustment will be made for this supportive analysis.

Another supportive analysis will include a multivariate Cox proportional hazards regression (CPH) method assessing independent predictors of OS after adjusting for potential confounders. Selection of covariates will primarily be based on *a priori* knowledge, and include age (continuous scale), prior cancer therapy (number of regimens will be entered on continuous scale), ECOG performance status (2 categories ECOG 0 and 1pooled together, and ECOG 2), and KRAS mutation status (coded as no=0, yes=1), and baseline body mass index (BMI). Covariates may be excluded from model if the number of patients representing 1 level of that variable is insufficient. A backward selection method will be used for variable selection; the p-value threshold will be set at 0.10. The proportional-hazard assumption will be tested by Schoenfeld residual analysis. The hazard ratio (HR) and 95% Wald confidence interval (CI) for the HR will be computed and presented in the results. The CPH analysis will be exploratory in nature.

8.4.2 Secondary Effectiveness Endpoints:

LBM: LBM change will be compared across treatment groups both by responder analysis and assessing absolute change.
Patients showing stabilization or improvement from baseline (≥0 kg change) will be considered LBM responders and those experiencing LBM reduction will be defined as non-responders. Patients not having a second evaluation for going off the study or any other reason will be considered non-responders. LBM response rate will be compared using Pearson chi-square test. Proportion by group, relative risk and odds ratio estimates will be presented with Cochran-Mantel-Haenszel (CMH) 95% confidence intervals (95% CI).

Absolute change will be compared between treatment groups using an analysis of covariance (ANCOVA) method (SAS GLM procedure) with treatment groups as the factor and baseline value as covariate. The difference in least-square means and 2-sided P values will be derived from the ANCOVA model.

EORTC: Longitudinal and descriptive analysis will be performed to assess patient reported outcomes (PRO). The EORTC survey results will be converted to domain specific scores using the prescribed algorithm (Appendix A). Scores on the EORTC QLQ-C30 functional and symptom scales will be summarized according to the randomization groups. The EORTC survey responses will be converted to domain specific scores using the prescribed algorithm. Scores on EORTC QLQ-C30 functional and symptom scales will be summarized according the randomization groups. Descriptive statistics of PRO score will be summarized by treatment group.

Patients showing stabilization or improvement from baseline (≤0 change) on any two of the three symptom scales (pain, fatigue, appetite loss) will be considered EORTC responders. Patients experiencing reduction or worsening on the symptom scales will be defined as non-responders. Patients missing the follow-up assessment will be considered non-responders. The comparison of responder rate and absolute change will be performed similar to as described for LBM.

Clinical Response Rate: Patients responding on LBM (≥0 kg change from baseline) and EORTC (≤0 change on any two three symptoms) will be considered Clinical Responders. Patients missing follow-up status will be considered non-responders. Clinical Response Rate will be compared using a Pearson chi-square test. Proportion by treatment groups and relative risk estimates will be presented with CMH 95% confidence intervals.

Platelet Count: Change in platelet count from baseline to follow-up will be compared between treatment groups using the ANCOVA method described above. The last visit cycle where at least 20 patients per group are on study will be used for follow-up data cut-off for this measure. Trend line plots displaying the median platelet count over time (visit cycles) by treatment groups will be generated.

Progression Free Survival (PFS): PFS will be compared by Kaplan-Meier method using log-rank test. Kaplan-Meier plots by treatment arm will be produced, and point estimates and 95% CIs will be summarized.

Objective Response Rate (ORR): The RR will be estimated by dividing the total number of patients with confirmed complete response (CR) and partial response (PR) by the total number of patients in the ITT population. Best overall response (BOR), determined from the sequence of cycle responses assessed, will be used for analysis purposes. Response rates will be compared between treatment groups using a Pearson Chi-square test and Wald 95% confidence intervals for the response rate in each treatment groups will be calculated.

Disease Control Rate (DCR): The DCR will be estimated by dividing the total number of confirmed CRs, PRs and stable disease (SD) by the total number of patients in the ITT population. DCR will be analyzed as described for ORR. Best overall response (BOR), will be used for DCR analysis. DCR will be analyzed as described for ORR.

Inferential analysis for the secondary endpoints, without multiplicity correction, will be carried out separately for stratification variables (age, prior therapy, ECOG performance status, KRAS mutation status, and prior regorafenib treatment status), One-way ANOVA and Fisher's exact tests will be used to compare differences across stratification groups. Proportions by treatment group along with 95% CI will be reported.

A two-tailed p value of <0.05 will be considered statistically significant for all secondary endpoints. SAS 9.2 or higher (SAS Institute Inc., Cary, NC) will be used for statistical analysis.

8.4.3 Additional Exploratory Analyses:

Additional analyses will be performed comparing the primary overall survival endpoint, the secondary endpoints, and incidence of SAE, between Clinical Responders and Non-responders. The statistical methodology in conducting these analyses will follow the same described in the primary and secondary efficacy analysis and safety analysis sections.

8.4.1 Pharmacodynamics (PK) analysis:

PK samples will be tested by XBiotech laboratory for a randomly selected patient population as defined in the study protocol. Descriptive statistics of PK parameters will be presented in a table format. The summary will include average MABp1 concentration at visit cycles along with the standard deviation and coefficient of variation (CV).

8.4.2 Analysis of Safety Outcomes:

The assessment of the safety profile will be based on adverse events, vital signs, physical examinations, and clinical laboratory measurements. All randomized subjects, will be included in the safety analysis.

Exposure duration will be summarized by treatment groups and presented in a table format. Exposure duration will be calculated as Last infusion date – C1D1 infusion date.

Adverse events will be grouped by System Organ Class (SOC) and Preferred Term (PT) within SOC according MedDRA coding dictionary (version 18.0). The number of subjects reporting at

least one adverse event by PT will be summarized across treatment groups. AEs will be summarized by SOC and PT, by relatedness to study treatment and age group.Tables summarizing PT by CTCAE grade and narratives for serious adverse events (SAE) will be presented for each treatment group. A listing of all adverse events by treatment group will be included in the appendix section.

The numbers and percentages of patients experiencing any adverse event (AE) and any serious adverse event (SAE) will be summarized per treatment group will be compared using Chi-square test. The numbers and percentages of patients experiencing any adverse events will also be summarized by maximum severity and relatedness to study treatment per treatment group. Clinical laboratory tests (hematology and chemistry) data will be listed for each subject. Laboratory data will be summarized by Xilonix™ and placebo group. Values outside the normal reference range will be flagged as high or low on the listings. Summary statistics for baseline, on-treatment and change from baseline values will be provided by treatment group in table format. Statistical process control charts showing baseline and post-baseline visit values in comparison with laboratory reference values will be provided for tests of interest.

Vital signs listings (include temperature, pulse, and blood pressure) for each subject will be provided in the appendix section of the Clinical Study Report. Summary statistics by treatment groups will be presented in tabular format.

9 List of Planned Tables

- Randomization by study site, stratified by country
- Subject disposition (all enrolled subjects) with reason for discontinuation, cumulative dose (for Xilonix™ group), and duration on study before discontinuation, broken down by treatment group
- Major protocol deviations (all enrolled subjects)
- Baseline demographic and clinical characteristics by treatment group (per-protocol population)
- Baseline demographic and clinical characteristics by treatment group (ITT population)
- Baseline laboratory data by treatment group (per-protocol population)
- Baseline laboratory data by treatment group (ITT population)
- Listing of patients excluded from efficacy analysis, with reason for exclusion
- Primary endpoint (overall survival) summary by treatment group (count and percent, median survival) (ITT and PP population)
- Secondary endpoints (LBM, platelet count, and EORTC score): summary of change by treatment group: baseline, follow-up, and change with confidence intervals (ITT population and PP population).
- SAE by treatment group (all randomized patients)
- Adverse events (counts, percent) by MedDRA preferred term by treatment group
- Detail listing of adverse events by treatment group (all randomized patients)
- Summary of vital signs and change from baseline over visits (all randomized patients)

- Exploratory summary on pharmacodynamics (PK) analysis findings (patients in Xilonix arm)
- Exploratory summary on Clinical Response Rate analysis findings (patients in Xilonix arm)

APPENDIX

ENGLISH

EORTC QLQ-C30 (version 3)

We are interested in some things about you and your health. Please answer all of the questions yourself by circling the number that best applies to you. There are no "right" or "wrong" answers. The information that you providewill remain strictly confidential.

Please fill in your initials: **aaaa**
Your birthdate (Day, Month, Year): **bdbdbccd**
Today's date (Day, Month, Year): 31 **bdbdbccd**

	Not at All	A Little	Quite a Bit	Very Much
1. Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase?	1	2	3	4
2. Do you have any trouble taking a long walk?	1	2	3	4
3. Do you have any trouble taking a short walk outside of the house?	1	2	3	4
4. Do you need to stay in bed or a chair during the day?	1	2	3	4
5. Do you need help with eating, dressing, washing yourself or using the toilet?	1	2	3	4

During the past week:	Not at All	A Little	Quite a Bit	Very Much
6. Were you limited in doing either your work or other daily activities?	1	2	3	4
7. Were you limited in pursuing your hobbies or other leisure time activities?	1	2	3	4
8. Were you short of breath?	1	2	3	4
9. Have you had pain?	1	2	3	4
10. Did you need to rest?	1	2	3	4
11. Have you had trouble sleeping?	1	2	3	4
12. Have you felt weak?	1	2	3	4
13. Have you lacked appetite?	1	2	3	4
14. Have you felt nauseated?	1	2	3	4
15. Have you vomited?	1	2	3	4
16. Have you been constipated?	1	2	3	4

Please go on to the next page

ENGLISH

During the past week:	Not at All	A Little	Quite a Bit	Very Much
17. Have you had diarrhea?	1	2	3	4
18. Were you tired?	1	2	3	4
19. Did pain interfere with your daily activities?	1	2	3	4
20. Have you had difficulty in concentrating on things, like reading a newspaper or watching television?	1	2	3	4
21. Did you feel tense?	1	2	3	4
22. Did you worry?	1	2	3	4
23. Did you feel irritable?	1	2	3	4
24. Did you feel depressed?	1	2	3	4
25. Have you had difficulty remembering things?	1	2	3	4
26. Has your physical condition or medical treatment interfered with your family life?	1	2	3	4
27. Has your physical condition or medical treatment interfered with your social activities?	1	2	3	4
28. Has your physical condition or medical treatment caused you financial difficulties?	1	2	3	4

For the following questions please circle the number between 1 and 7 that best applies to you

29. How would you rate your overall health during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

30. How would you rate your overall quality of life during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

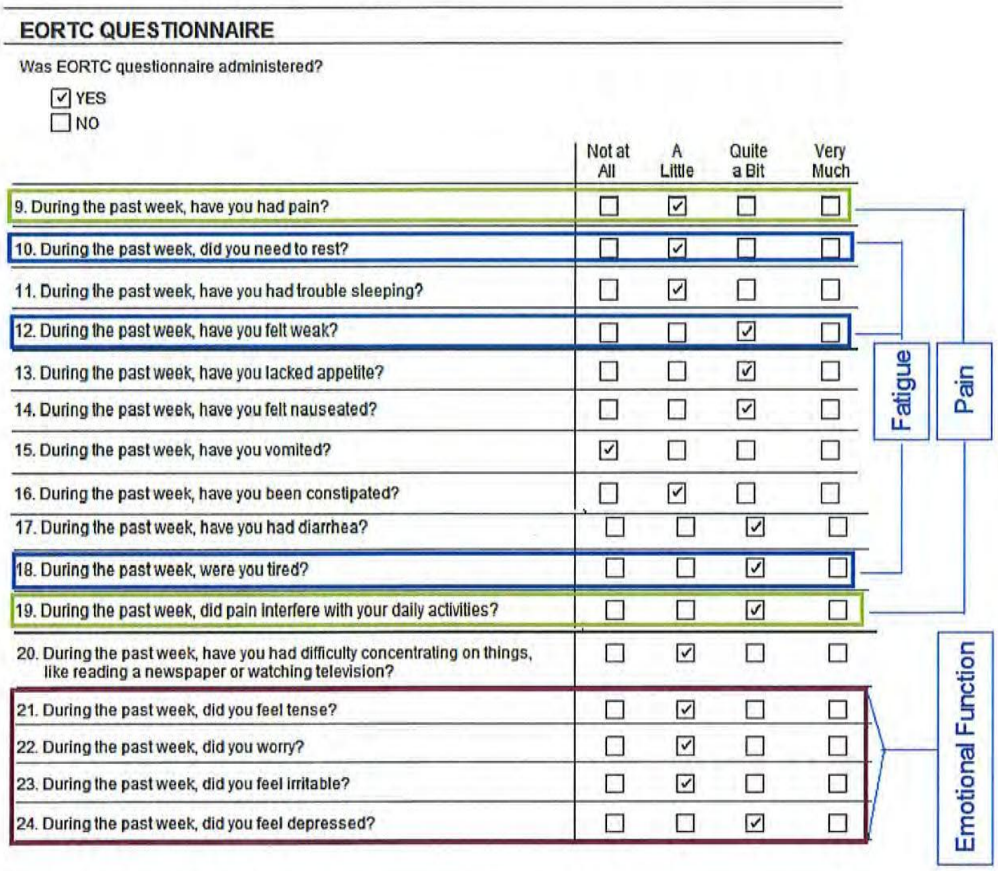
EORTC QLQ-30 Scoring Procedure
The European Organization for Research and Treatment of Cancer (EORTC) developed an integrated modular approach for evaluating quality of life (QoL) in patients participating in cancer clinical trials. The EORTC QLQ-C30 (version 3.0) questionnaire consists of three scales; functional scale, symptom scale, and a global QoL scale.

Inclusion criteria no 2 of the study protocol requires that the subjects need to show "Evidence of reduced function or presence of cancer related symptoms as determined by EORTC QLQ-C30". In order to determine eligibility, it is necessary to calculate the functional and symptom scores at the screening visit. The following steps describe the scoring procedure and the algorithm for computing the item level score (1). Please refer to table A1 for a list of questions that are included in the specific functional and symptom scales.

Calculate Raw Score (RS):
There are four response categories for each item (question). Based on the response to a given question, please assign a numeric value to the individual item; i.e. Not at All =1, A Little =2, Quite a Bit =3, and Very Much =4.
Then calculate raw score for a particular scale by taking the average of the individual item scores $[(I_1+I_2+I_3...I_n)/n]$.
For example, Fatigue is composed of three items, question 10, 12, and 18. So the raw score will be (response to Q10+response to Q12+ response to Q18)/3.

Linear Transformation:
All the scales need to be transformed to a linear scale and the final score should range from 0 to 100. Use the following formula to apply linear transformation and obtain the final score (S).
S= {1-((RS-1)/range)} X 100
Range: range is the difference between maximum and minimum possible scores. With 4 response categories in each item, the maximum and minimum scores can take a value of 4 to 1. Hence the **Range will be 3 for all the functional and symptom scales**.

Screenshot showing a sample EORTC QLQ-C30 questionnaire from the CRF page.



EORTC QUESTIONNAIRE

Was EORTC questionnaire administered?

☑ YES
☐ NO

	Not at All	A Little	Quite a Bit	Very Much
9. During the past week, have you had pain?	☐	☑	☐	☐
10. During the past week, did you need to rest?	☐	☑	☐	☐
11. During the past week, have you had trouble sleeping?	☐	☑	☐	☐
12. During the past week, have you felt weak?	☐	☐	☑	☐
13. During the past week, have you lacked appetite?	☐	☐	☑	☐
14. During the past week, have you felt nauseated?	☐	☐	☑	☐
15. During the past week, have you vomited?	☑	☐	☐	☐
16. During the past week, have you been constipated?	☐	☑	☐	☐
17. During the past week, have you had diarrhea?	☐	☐	☑	☐
18. During the past week, were you tired?	☐	☐	☑	☐
19. During the past week, did pain interfere with your daily activities?	☐	☐	☑	☐
20. During the past week, have you had difficulty concentrating on things, like reading a newspaper or watching television?	☐	☑	☐	☐
21. During the past week, did you feel tense?	☐	☑	☐	☐
22. During the past week, did you worry?	☐	☑	☐	☐
23. During the past week, did you feel irritable?	☐	☑	☐	☐
24. During the past week, did you feel depressed?	☐	☐	☑	☐

Fatigue
Pain
Emotional Function

Examples 1, Emotional functioning:

For example, the "Emotional functioning" (EF) scale is calculated from items Q21, Q22, Q23, Q24. It is scaled to range from 0 to 100, with a high score indicating high emotional functioning. As presented in the sample questionnaire (figure A1), the score for Q21 to Q24 are 2, 2, 2, and 3 respectively.

Raw Score (RS) = average of 2, 2, 2, and 3; which is (2+2+2+3)/4= 2.25

Linear Score (S) = {1-((RS-1)/range)} X 100 = {1-((2.25-1)/3)} X 100 = {1-((1.25)/3)} X 100

Which is (1-0.42) X 100 = 0.583X100

The final linear score for EF = 58.3, which meets the inclusion criteria (≤89).

Example 2, Fatigue:

Fatigue scale includes items Q10, Q12, and Q18; for which the individual scores are 2 for Q10, 3 for Q12, and 3 for Q18.

Raw Score (RS) = average of 2, 3, and 3;

= (2+3+3)/3= 2.67

Linear Score (S) = {((RS-1)/range)} X 100 = {((2.67-1)/3)} X 100 = 55.6

With a base fatigue score of 55.6 the subject meets the inclusion criteria of >10.

Example 3, Pain:

The scale for Pain includes two items Q9, Q19; for which the individual scores in figure 1A are 2 for Q9, 3 for Q19.

Raw Score (RS) = (2+3)/2= 2.5

Linear Score (S) = {((RS-1)/range)} X 100 = {((2.5-1)/3)} X 100 = 50.0

With a baseline symptom score of 50.0 for pain, the subject meets the inclusion criteria of >10.

Reference

1: Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. Journal of the National Cancer Institute. 85: 365-376, 1993.

2: John Tukey, Exploratory Data Analysis, Addison-Wesley, 1977, pp. 43-44.

Table Shells

1) Patient Disposition:

	Treatment Group		
	Xilonix+BSC (N= XXX) n (%)	Placebo+BSC N= XXX) n (%)	Total N=XXX
Randomized	XXX	XXX	XXX
Intent to Treat Population (ITT)	XXX (XXX)	XXX (XXX)	XXX (XXX)
Did not receive study drug	XX (X)	XX (X)	XX (X)
Discontinued from study	XX (XX)	XX (XX)	XX (XX)
Reason for discontinuation			
Adverse Event	XX (XX)	XX (XX)	XX (XX)
Death	XX (XX)	XX (XX)	XX (XX)
Lost to Follow-Up	XX (XX)	XX (XX)	XX (XX)
Other†	XX (XX)	XX (XX)	XX (XX)
Removed from Study by PI†	XX (XX)	XX (XX)	XX (XX)
Subject Withdrew Consent	XX (XX)	XX (XX)	XX (XX)

†Listing of specific reason will be provided

2) Important Protocol Deviations:

Type of Protocol Deviation	Xilonix+BSC (n=XXX) n (%)	Placebo+BSC (XXX) n (%)
Deviation Category, n(%)	XX (XX)	XX (XX)

3) Enrollment at Clinical Site by Country:

Country	Site Number	Number of Patients Enrolled
XXXX	XXXX	X
	XXXX	X

4) Patient Demographic and Baseline Characteristics

	Treatment Group		Total (N= XXX)
	Xilonix+BSC (N= XXX)	Placebo+BSC (N=XXX)	
Age, year			
Mean	XX±XX	XX±XX	XX±XX
Median	XX	XX	XX
Min-Max	XX-XX	XX-XX	XX-XX
Age distribution, n(%)			
<65 years	XXX (XX%)	XXX (XX%)	XXX (XX%)
≥65 to <75 years	XXX (XX%)	XXX (XX%)	XXX (XX%)
>75 years	XXX (XX%)	XXX (XX%)	XXX (XX%)
Sex, n(%)			
Female	XX (XX)	XX (XX)	XX (XX))
*Race, n(%)			
White	XX (XX)	XX (XX)	XX (XX)
Black	XX (XX)	XX (XX)	XX (XX)
Asian	XX (XX)	XX (XX)	XX (XX)
Geographic Region , n(%)			
XXXX	XX (XX)	XX (XX)	XX (XX)
XXXX	XX (XX)	XX (XX)	XX (XX)
*KRAS Mutation Status, n(%)			
KRAS Mutation	XX (XX)	XX (XX)	XX (XX)
KRAS wild-type	XX (XX)	XX (XX)	XX (XX)
Test Not Done	XX (XX)	XX (XX)	XX (XX)
ECOG Performance Status			
0	XX (XX)	XX (XX)	XX (XX)
1	XX (XX)	XX (XX)	XX (XX)
2	XX (XX)	XX (XX)	XX (XX)
Baseline Weight, kg			
Mean	XX±XX	XX±XX	XX±XX
Median	XX	XX	XX
Min-Max	XX-XX	XX-XX	XX-XX
Histology, n(%)			
Adenocarcinoma	XX (XX)	XX (XX)	XX (XX)
Adenocarcinoma in situ	XX (XX)	XX (XX)	XX (XX)
Other	XX (XX)	XX (XX)	XX (XX)
Prior Therapies, n(%)			

	Treatment Group		Total (N= XXX)
	Xilonix+BSC (N= XXX)	Placebo+BSC (N=XXX)	
XXX	XX (XX)	XX (XX)	XX (XX)
XXX	XX (XX)	XX (XX)	XX (XX)
Number of prior chemotherapy regimens, n(%)			
2	XX (XX)	XX (XX)	XX (XX)
3	XX (XX)	XX (XX)	XX (XX)
4	XX (XX)	XX (XX)	XX (XX)
5	XX (XX)	XX (XX)	XX (XX)
≥6	XX (XX)	XX (XX)	XX (XX)

5) Exposure by Treatment Groups

	Xilonix+BSC (N=XXX)			Placebo+BSC (N=XXX)		
	N	Mean±SD	Median (IQR)	N	Mean±SD	Median (IQR)
All Subjects	XXX	XX±XX	XX-XX	XXX	XX±XX	XX-XX
Female	XXX	XX±XX	XX-XX	XXX	XX±XX	XX-XX
Male	XXX	XX±XX	XX-XX	XXX	XX±XX	XX-XX
Age group=<65yr	XXX	XX±XX	XX-XX	XXX	XX±XX	XX-XX
Age group=65-75yr	XXX	XX±XX	XX-XX	XXX	XX±XX	XX-XX
Age group=>75yr	XXX	XX±XX	XX-XX	XXX	XX±XX	XX-XX

6) Clinical Response Rate by Treatment Groups ITT Population:

	Xilonix+BSC	Placebo+BSC
N	XXX	XXX
Clinical Response, n (%, 95% CI)	XX (XX, XX-XX)	XX (XX, XX-XX)
Difference (effect size)	XX%	
P value from Pearson Chi-Square test (one-tailed)	0.01	
Relative Risk (95% CI)	X.XX (X.XX to X.XX)	

7) Clinical Response Rate by Treatment Groups PP Population:

	Xilonix+BSC	Placebo+BSC
N	XXX	XXX
Clinical Response, n (%, 95% CI)	XX (XX, XX-XX)	XX (XX, XX-XX)
Difference (effect size)	XX%	
P value from Pearson Chi-Square test (one-tailed)	0.01	
Relative Risk (95% CI)	X.XX (X.XX to X.XX)	

8) Objective Response Rate by Treatment Groups ITT Population:

	Xilonix+BSC	Placebo+BSC
N	XXX	XXX
Objective Response, n (%, 95% CI)	XX (XX, XX-XX)	XX (XX, XX-XX)
Difference (effect size)	XX%	
P value from Pearson Chi-Square test (one-tailed)	0.01	
Relative Risk (95% CI)	X.XX (X.XX to X.XX)	

9) Objective Response Rate by Treatment Groups PP Population:

	Xilonix+BSC	Placebo+BSC
N	XXX	XXX
Objective Response, n (%, 95% CI)	XX (XX, XX-XX)	XX (XX, XX-XX)
Difference (effect size)	XX%	
P value from Pearson Chi-Square test (one-tailed)	0.01	
Relative Risk (95% CI)	X.XX (X.XX to X.XX)	

10) Disease Control Rate by Treatment Groups ITT Population:

	Xilonix+BSC	Placebo+BSC

N	XXX	XXX
Disease Control Rate, n (%, 95% CI)	XX (XX, XX-XX)	XX (XX, XX-XX)
Difference (effect size)	XX%	
P value from Pearson Chi-Square test (one-tailed)	0.01	
Relative Risk (95% CI)	X.XX (X.XX to X.XX)	

11) Disease Control Rate by Treatment Groups PP Population:

	Xilonix+BSC	Placebo+BSC
N	XXX	XXX
Disease Control Rate, n (%, 95% CI)	XX (XX, XX-XX)	XX (XX, XX-XX)
Difference (effect size)	XX%	
P value from Pearson Chi-Square test (one-tailed)	0.01	
Relative Risk (95% CI)	X.XX (X.XX to X.XX)	

12) Secondary Endpoints Results: Change from Baseline

	Placebo+BSC (n= xxx)			Xilonix+BSC (n= xxx)			P*
	Baseline Mean (95% CI)	Follow-up Mean (95% CI)	Change LS Mean±SE	Baseline Mean (95% CI)	Follow-up Mean (95% CI)	Change LS Mean±SE	
Platelet, 1000/cu mm	XXX±XXX	XXX±XXX	X.XX±X.XX	XXX±XXX	XXX±XXX	X.XX±X.XX	0.XXX
PRO: Global QOL Score	XXX±XXX	XXX±XXX	X.XX±X.XX	XXX±XXX	XXX±XXX	X.XX±X.XX	0.XXX
PRO Symptom Score: Pain	XXX±XXX	XXX±XXX	X.XX±X.XX	XXX±XXX	XXX±XXX	X.XX±X.XX	0.XXX
PRO Symptom Score: Fatigue	XXX±XXX	XXX±XXX	X.XX±X.XX	XXX±XXX	XXX±XXX	X.XX±X.XX	0.XXX
PRO Symptom Score: Appetite	XXX±XXX	XXX±XXX	X.XX±X.XX	XXX±XXX	XXX±XXX	X.XX±X.XX	0.XXX
Lean Body Mass, kg	XXX±XXX	XXX±XXX	X.XX±X.XX	XXX±XXX	XXX±XXX	X.XX±X.XX	0.XXX

13) All Adverse Events by CTCAE Grade

XBiotech USA, Inc., Protocol 2012-PT023 V3.6 Interim Analysis #X, Data Date: ddmmmyyyy

Body System or Organ Class	Group: Xilonix+BCS (N=XXX) Statistics n (%)					Group: Placebo+BCS (N= XX) Statistics n (%)				
AE Preferred Term	Grade I/II	Grade III	Grade IV	Grade V	Total	Grade I/II	Grade III	Grade IV	Grade V	Total
XXXXXX	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)
XXXXXX	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)
XXXXXX	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)
XXXXXX	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)

The above table will be provided for SAEs.

14) Drug-related AEs

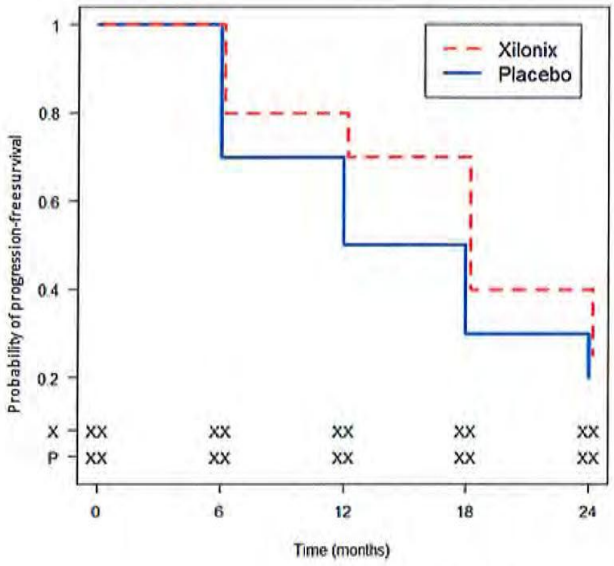
Body System or Organ Class	Group: Xilonix+BCS; Related (N=XX) Statistics n (%)					Group: Xilonix+BCS; Not Related (N=XX) Statistics n (%)				
AE Preferred Term	Grade I/II	Grade III	Grade IV	Grade V	Total	Grade I/II	Grade III	Grade IV	Grade V	Total
XXXXXX	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)
XXXXXX	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)

15) AEs by Age Group:

Body System or Organ Class	Group: Xilonix+BCS; Under 65 year; (N=88) Statistics n (%)					Group: Xilonix+BCS; 65 to 75 year; (N=56) Statistics n (%)					Group: Xilonix+BCS; Over 75 year; (N=15) Statistics n (%)				
AE Preferred Term	Grade I/II	Grade III	Grade IV	Grade V	Total	Grade I/II	Grade III	Grade IV	Grade V	Total	Grade I/II	Grade III	Grade IV	Grade V	Total
XXXXXX	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)
XXXX	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)	XX (XX.X%)

16) Table and Figure for Progression-free Survival:

Kaplan-Meier Curves Comparing Progression-free Survival

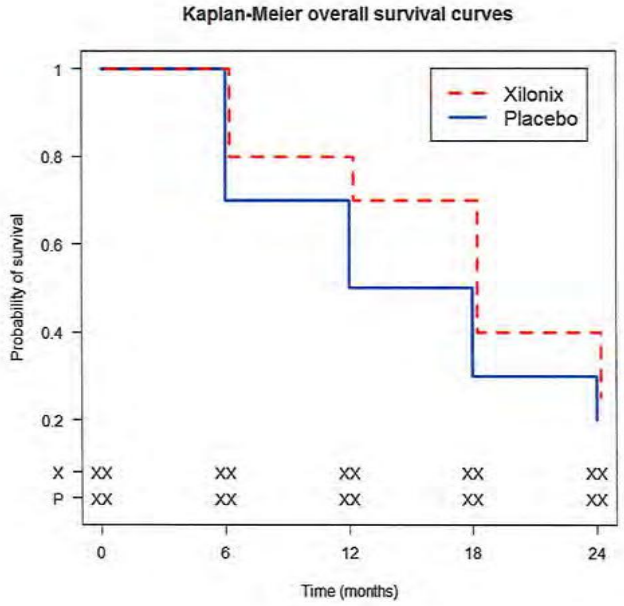


Summary Table of Kaplan-Meier Log-Rank Test

Stratum	Total subjects	Failed	Censored	Percent Censored	25th %ile 95% CI	50th %ile 95% CI	75th %ile 95% CI
Xilonix	XX	XX	XX	XX.X%	XX.X (X.X, X.X)	XX.X (X.X, X.X)	XX.X (X.X, X.X)
Placebo	XX	XX	XX	XX.X%	XX.X (X.X, X.X)	XX.X (X.X, X.X)	XX.X (X.X, X.X)
Log-rank test		X^2=XX (SE= XX)	P=0.XXX				

TABLES AND FIGURES FOR OVERALL SURVIVAL ANALYSIS

Summary Figure of Kaplan-Meier Curves of Overall Survival by Treatment Group



Summary Table of Kaplan-Meier Event Summary, by Treatment Group

	Xilonix (N=XX)				*Placebo (N=XX)*			
Survival Time (months)	Survival Prob	Survival Standard Error	Number Failed	Number Remaining on Therapy	Survival Prob	Survival Standard Error	Number Failed	Number Remaining on Therapy
0.0	0.XXX	0.XXX	XX	XX	0.XXX	0.XXX	XX	XX
6.0	0.XXX	0.XXX	XX	XX	0.XXX	0.XXX	XX	XX
12.0	0.XXX	0.XXX	XX	XX	0.XXX	0.XXX	XX	XX
18.0	0.XXX	0.XXX	XX	XX	0.XXX	0.XXX	XX	XX

Summary Table of Kaplan-Meier Log-Rank Test

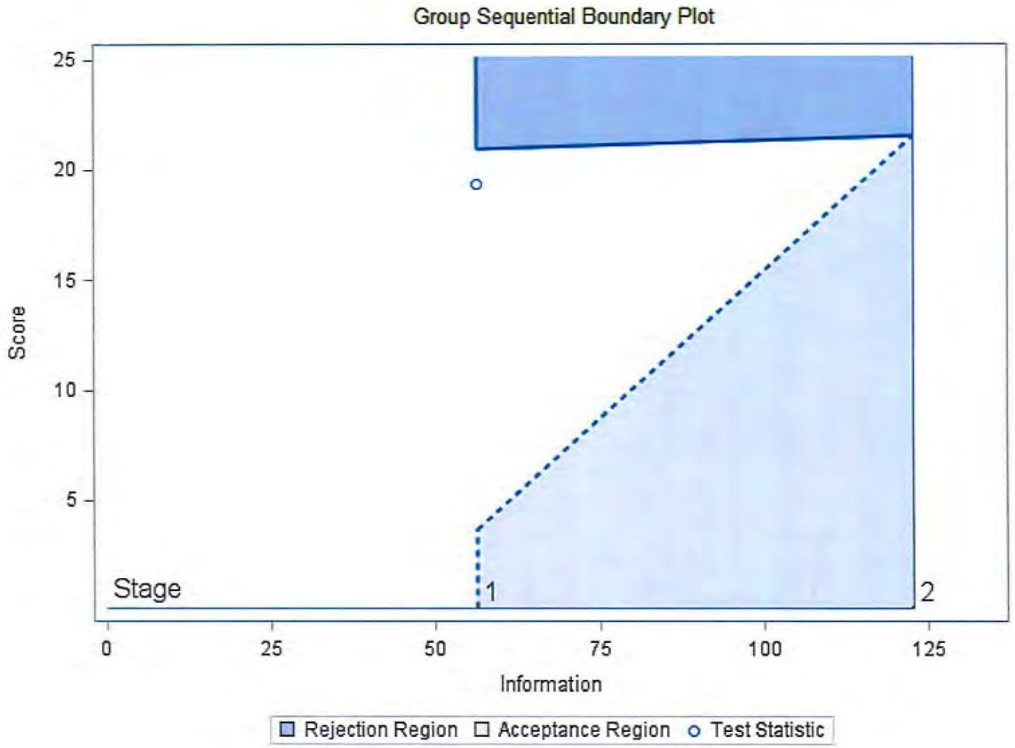
Stratum	Total subjects	Failed	Censored	Percent Censored	25th %ile 95% CI	50th %ile 95% CI	75th %ile 95% CI
Xilonix	XX	XX	XX	XX.X%	XX.X (X.X, X.X)	XX.X (X.X, X.X)	XX.X (X.X, X.X)
Placebo	XX	XX	XX	XX.X%	XX.X (X.X, X.X)	XX.X (X.X, X.X)	XX.X (X.X, X.X)
Log-rank test		X^2=XX (SE= XX)	P=0.XXX				

Summary Table of Group Sequential Boundary and Interim Analysis Results

Stage	Information Level		Score Scale			p-value scale			Action*
	Proportion	Actual	Futility Boundary	Efficacy Boundary	Estimate	Futility Boundary	Efficacy Boundary	Estimate	
1	50%	XX.X%	X.XX	X.XX	X.XX	0.XXXX	0.XXXX	0.XXXX	XXXXXX
2	75%	XX.X%	X.XX	X.XX	X.XX	0.XXXX	0.XXXX	0.XXXX	XXXXXX
3	100%	XX.X%	X.XX	X.XX	X.XX	0.XXXX	0.XXXX	0.XXXX	XXXXXX

*Continue; Stop for Futility; or Stop for Efficacy

Summary Figure of Group Sequential Boundary



Listing of Subject Disposition / Survival

Subject ID	Sex	Age (years)	Treatment Group	Status*	Discontinuation Reason	Time to Status
1	M/F	X.XX	X.XX	XXXXXXXXXX	XXXXXXXXXX	XXXXXX
2	M/F	X.XX	X.XX	XXXXXXXXXX	XXXXXXXXXX	XXXXXX
3	M/F	X.XX	X.XX	XXXXXXXXXX	XXXXXXXXXX	XXXXXX

*Never Dosed; On Study; Discontinued - Death; Discontinued - Other

Programming Notes:

Please sort by Subject ID

For Status = Discontinued, Please add in if the subject discontinued due to death, or for any other reason