



## Table of Contents

1.	Introduction.....	6
2.	Study Objectives .....	7
3.	Study Design.....	8
3.1	Design Overview .....	8
3.2	Determination of Sample Size.....	8
4.	General Statistical Considerations .....	9
4.1	General Principles.....	9
4.2	Handling of Dropouts and Screen Failures.....	9
4.3	Handling of Missing Data.....	10
4.3.1	Imputation rules.....	10
4.4	Interim Analyses and Data Monitoring .....	10
4.5	Data Rules.....	10
4.6	Validity Review and Protocol Deviations .....	10
5.	Analysis Sets.....	11
6.	Statistical Methodology .....	11
6.1	Population characteristics .....	11
6.1.1	Disposition .....	11
6.1.2	Demographics and Baseline Characteristics .....	11
6.1.3	Medical history.....	12
6.1.4	Drug Exposure and Compliance .....	13
6.1.5	Prior and Concomitant Medications.....	13
6.2	Efficacy.....	14
6.2.1	Primary Efficacy Analyses.....	14
6.2.2	Secondary Efficacy Analyses.....	15
6.2.3	Exploratory Efficacy Analysis .....	18
6.3	Pharmacokinetics / pharmacodynamics.....	21
6.4	Safety.....	21
6.4.1	Adverse events .....	21
6.4.2	Deaths.....	23
6.4.3	Clinical Laboratory Data.....	23
6.4.4	Other Safety Analyses.....	24
6.5	Maximum tolerated dose (MTD)/Recommended phase 2 dose (RP2D).....	24
6.6	Analysis of Other Endpoints.....	24
6.6.1	ECOG Performance Status.....	24
6.6.2	Brief Pain Index – Short Form (BPI-SF) .....	25
6.6.3	Resource Utilization Questionnaire .....	25
6.6.4	Long Term Safety.....	25
6.7	Examination of Subgroups .....	25
7.	Document history and changes in the planned statistical analysis .....	25
8.	References.....	26

## Abbreviations

AE	adverse event
BMI	body mass index
BOR	bortezomib
CI	confidence interval
CR	complete response
CRF	case report form
CT	computed tomography
ctDNA	circulating tumor DNA
CTCs	circulating tumor cells
CTCAE	Common Terminology Criteria for Adverse Events
DEM	dose escalation meeting
DES	dose evaluable set
DEX	dexamethasone
DK	decay correction factor
DL	dose level
DLT	dose limiting toxicity
DNA	deoxyribonucleic acid
DOR	duration of response
EBRT	external beam radiotherapy
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EDC	electronic data capture
EOS	end of study
EOT	end of treatment
EU	European Union
FDA	United States Food and Drug Administration
Hb	hemoglobin
HR	hazard ratio
IC	investigator's choice
ICF	informed consent form
ICH	International Conference on Harmonization
ICTP	Type I collagen telopeptide
ID	identification number

IDMC	Independent Data Monitoring Committee
IEC	Independent Ethics Committee
IF	immunofixation
IMP	investigational medicinal product
IMWG	International Myeloma Working Group
INR	international normalized ratio (reagent-independent prothrombin ratio)
IRB	Institutional Review Board
ISS	International Staging System for Multiple Myeloma
ITT	intent-to-treat
IV	intravenous
IxRS	Interactive Voice/Web Response System
kBq	kiloBecquerel; SI unit of radioactivity
kg	kilogram
LLOQ	lower limit of quantification
LDH	lactate dehydrogenase
MedDRA	Medical Dictionary for Regulatory Activities
mg	milligram
mL	milliliter
MRI	magnetic resonance imaging
MTD	maximum tolerated dose
NCI CTCAE	National Cancer Institute Common Toxicity Criteria for Adverse Events
ORR	objective response rate
OR	overall response
OS	overall survival
PCR	polymerase chain reaction
PD	progressive disease
PEP	protein electrophoresis
PET-CT	positron emission tomography – computed tomography
PFS	progression-free survival
PK	pharmacokinetic
PR	partial response
PR	PR interval in the ECG
PS	performance status
PSA	prostate-specific antigen
PT(-INR)	prothrombin time(-international normalized ratio)

PTT	partial thromboplastin time
QT	QT interval in ECG
RAVE	Medidata Rave; electronic data capture tool
RBC	red blood cell
RNA	ribonucleic acid
RP2D	recommended phase 2 dose
RUQ	resource utilization questionnaire
SAE	serious adverse event
SAF	safety analysis set
SAP	statistical analysis plan
SC	subcutaneous
sCR	stringent complete response
SD	stable disease
SRE	skeletal-related event
SSE	symptomatic skeletal event
SUSAR	suspected unexpected serious adverse reaction
TEAE	treatment-emergent adverse event
TTP	time to progression
ULN	upper limit of normal
US	United States
VGPR	very good partial response
WBC	white blood cell
WHO DD	World Health Organization Drug Dictionary
WPS	worst pain score
wk	week

## 1. Introduction

Protocol 18987 is a phase 1b/2 trial to evaluate the safety of radium-223 dichloride (BAY 88-8823) in combination with bortezomib (BOR) and dexamethasone (DEX) in patients with relapsed multiple myeloma. The study has two parts: Phase 1b Dose escalation and Phase 2 double blinded randomization.

Two cohorts at 2 different dose levels are planned for phase 1b dose escalation part. After maximum tolerated dose/recommended phase 2 dose is identified from dose escalation cohorts, the blinded randomized phase 2 part of the study will start and the efficacy and safety of radium-223 dichloride in combination with BOR and DEX versus placebo in combination with BOR and DEX will be evaluated.

The following analyses are planned in this study:

- The final analysis of phase 1b data will be conducted when all subjects in the phase 1b part have completed the end of treatment visit (including the maintenance therapy), or have discontinued early.
- For the phase 2 part, the primary analysis will be conducted after all subjects have completed the end of radum-223 dichloride visit and the final analysis will be conducted when all subjects enrolled have completed the end of treatment visit (also including the maintenance therapy).

The purpose of this statistical analysis plan (SAP) is to describe the planned analysis and data presentations for this study. This SAP is written in accordance with the principles in the International Conference on Harmonisation (ICH) guidances ICH-E3 (Good Clinical Principles), ICH-E9 (Statistical Principles for Clinical Trials), and the Bayer standard operating procedure BDP-SOP-060.

The Statistical Analysis Plan (SAP) was written based on the following documentation:

<b>Document</b>	<b>Date</b>	<b>Version</b>
Protocol	18 May 2018	3.0
Annotated eCRF - Subject Case Report Forms	31 January 2017	3.0 PROD
Xofigo Biostatistical Project Plan	7 December 2016	2.0
Xofigo Biostatistical Project Plan_v2_07Dec2016 final_addendum.docx	May 17, 2018	

## 2. Study Objectives

### Phase 1b part (open-label)

Primary objectives:

- To evaluate the safety of the combination of radium-223 dichloride plus bortezomib (BOR) and dexamethasone (DEX)
- To determine the dose of radium-223 dichloride that will be used in the phase 2 part of the study (maximum tolerated dose [MTD] or recommended phase 2 dose [RP2D])

Secondary objectives:

- To evaluate the combined complete response (CR) + very good partial response (VGPR), as determined by International Myeloma Working Group (IMWG) uniform response criteria

### Phase 2 part (double-blind, randomized)

Primary objectives:

- To compare radium-223 dichloride versus placebo in addition to background treatment with BOR plus DEX in terms of combined CR+VGPR rate in relapsed multiple myeloma subjects, as determined by International Myeloma Working Group (IMWG) uniform response criteria.

Secondary objectives:

- To evaluate safety
- To evaluate the objective response rate (ORR), as determined by IMWG uniform response criteria
- To evaluate overall survival (OS)
- To evaluate progression-free survival (PFS)
- To evaluate duration of response (DOR)

Exploratory objectives:

- To evaluate time to first on-study symptomatic skeletal event (SSE)
- To evaluate bone biomarkers
- To evaluate change in bone lesions
- To evaluate change in MM disease markers
- To evaluate circulating tumor DNA (ctDNA) and circulating tumor cells (CTCs)
- To evaluate time to opioid use for cancer pain in subjects without opioid use at baseline
- To evaluate time to pain progression

### 3. Study Design

#### 3.1 Design Overview

This study will be conducted in 2 parts. The phase 1b part will be an international, phase 1b, open-label, dose-escalation assessment of radium-223 dichloride administered with BOR and DEX in subjects with relapsed MM. The phase 2 part will be an international, phase 2, double-blind, randomized, placebo-controlled assessment of radium-223 dichloride versus placebo administered with BOR and DEX, in subjects with relapsed MM.

For the **Phase 1b** open-label part of the study, 2 dose levels of radium-223 dichloride were administered, following sequential dose escalation using a '3 + 3' design, in order to determine maximum tolerated dose/recommended phase 2 dose (MTD/RP2D) for phase 2:

- **Cohort 1:** 33 kiloBecquerel (kBq)/kg body weight every 6 weeks for a total of 6 radium-223 dichloride doses in combination with BOR and DEX
- **Cohort 2:** 55 kBq/kg body weight every 6 weeks for a total of 6 radium-223 dichloride doses in combination with BOR and DEX

For the **Phase 2** double-blind randomized part of the study, the MTD/RP2D, as determined by the phase 1b part of the study, will be administered once every 6 weeks for a total of 6 radium-223 dichloride doses.

Up to approximately total 12 subjects in the 2 dose cohorts combined will be treated in the phase 1b part of the study and approximately 100 subjects will be enrolled in the phase 2 part of the study.

#### 3.2 Determination of Sample Size

##### Phase 1b

To evaluate the MTD/RP2D for radium-223 dichloride in combination with BOR/DEX in subjects with relapsed multiple myeloma, two cohorts with up to 12 subjects are planned for dose escalation part. The sample size was based on feasibility, not formal statistical calculation. Subjects not evaluable for assessment of DLT for dose escalation part due to early discontinuation for reasons other than DLT will be replaced until the MTD/RP2D has been determined. The number of subjects is considered to be sufficient for a 3 + 3 dose escalation design to determine whether to start the next dose level, using DLT definitions and the 3 + 3 dose escalation rules.

##### Phase 2



The primary endpoint is combined CR+VGPR rate in phase 2 part of the study. Approximately 100 subjects will be randomized at a 1:1 ratio to two treatment arms: radium-223 dichloride + BOR/DEX (study arm), and placebo + BOR/DEX (control arm). The placebo + BOR/DEX arm is anticipated to have combined CR+VGPR rate of 35%. The radium-223 dichloride + BOR/DEX arm will be considered promising if true combined CR+VGPR rate will be 55% or higher. Assuming one-sided alpha of 0.2 and 50 randomized subjects in each arm, there will be a 87% power to detect a treatment difference of 0.2 in combined CR+VGPR rate between two arms.

## **4. General Statistical Considerations**

### **4.1 General Principles**

All data analyses will be carried out using the software package SAS<sup>®</sup> Version 9.2 or higher (SAS Institute Inc., Cary, NC, USA).

For continuous endpoints, the mean, the standard deviation, median, minimum, and maximum will be provided. For categorical data, the frequency and percent distributions will be provided.

In general, all analyses will be performed for phase 1b and phase 2, separately. All data will be listed, and descriptive summary tables, including demographics and baseline characteristics, will be provided by dose level cohort and overall for phase 1b, and by treatment arm and overall for phase 2. Any test of comparison performed for phase 2 will be one-sided at 20% level of significance unless otherwise specified.

If more than one value is reported at a scheduled visit, the latest value will be used.

### **4.2 Handling of Dropouts and Screen Failures**

A “dropout” is defined as a subject who has received the first dose of study treatment and discontinues study participation prematurely for any reason. For phase 1b part, subjects not evaluable for assessment of DLT due to early discontinuation for reasons other than DLT or subjects who have not completed 2 doses of radium-223 dichloride and 2 cycles of bortezomib will be replaced until the MTD/RP2D has been determined. For phase 2 part, a subject who prematurely discontinues the study will be counted toward the total enrollment goal and will not be replaced.

A subject who, for any reason (e.g., failure to satisfy the selection criteria) terminates the study before the first dose of radium-223 dichloride for phase 1b or before randomization for phase 2 is regarded as a “screening failure”.

### 4.3 Handling of Missing Data

In order to achieve the goal of a well conducted clinical trial according to ICH Good Clinical Practice (ICH-GCP), every effort should be made to collect all data. However, despite best efforts, it may be inevitable that missing or incomplete data are reported. All missing or partial data will be presented in the subject data listing, as they are recorded on the Case Report Form (CRF). Except as noted, missing data will not be imputed or carried forward in any statistical analysis.

#### 4.3.1 Imputation rules

Unless otherwise specified data will not be imputed, and rules for imputation are specified in the dataset specification documents following rules described in Xofigo Biostatistical Project Plan [3].

### 4.4 Interim Analyses and Data Monitoring

No formal interim analysis will be performed.

In the event that enrollment is halted by the Steering Committee after the review of safety data, an interim analysis may be performed on all evaluable subjects during the phase 1.

In the event that enrollment is halted by the IDMC after the review of safety data, an interim analysis may be performed on all randomized subjects during the phase 2.

### 4.5 Data Rules

Baseline value for demographic and baseline characteristic table, as well as that for efficacy and other endpoints is defined as the last available (non-missing) value on or before the date of randomization for phase 2 part.

Baseline value for demographic and baseline characteristic table, as well as that for efficacy and safety endpoints in the phase 1b part and safety endpoints in the phase 2 parts is defined as the last available (non-missing) value on or before the first dose date of any study drug.

If more than one value is reported at a scheduled visit, the last valid (non-missing) value will be used in the by visit summaries. All values will be included in the listings.

### 4.6 Validity Review and Protocol Deviations

The results of validity (protocol deviations) review meetings will be documented in the Validity Review Reports and may comprise decisions and details relevant for statistical evaluation. Any changes to the statistical analysis prompted by the results of the validity review meeting will be documented in an amendment or, if applicable, in a supplement to this SAP.

## 5. Analysis Sets

**Phase 1b Safety Set (SAF):** All subjects who received at least one dose of study treatment. Subjects will be included in the analyses according to the treatment they actually received.

**Phase 1b Dose Evaluable Set (DES):** Subjects in dose escalation cohorts who meet the minimum drug exposure criterion (at least 2 doses of radium-223 dichloride plus a follow-up of 21 days after the second radium-223 dichloride dose) or experience a DLT (during the DLT observation window). Subjects not evaluable for assessment of DLT due to early discontinuation for reasons other than DLT or subjects who have not completed 2 doses of radium-223 dichloride and 2 cycles of BOR will be replaced and will not be included in the DES.

**Phase 2 Intent-to-treat (ITT):** All randomized subjects in phase 2. The phase 2 ITT population will be the primary analysis set for all efficacy endpoints. Subjects will be included in the analyses according to the treatment to which they are randomized. The definition of the ITT population is identical to that of the FAS (Full Analysis Set) population.

**Phase 2 Safety Set (SAF):** All subjects who received at least one dose of study treatment. Subjects will be included in the analyses according to the treatment they actually received. Safety population will be used in the analyses of all safety endpoints.

## 6. Statistical Methodology

### 6.1 Population characteristics

#### 6.1.1 Disposition

For the phase 1b part of the study, the number and percentage of subjects screened, and treated will be summarized by cohort and overall. For the phase 2 part of the study, the number and percentage of subjects screened, randomized, and treated will be presented by treatment group and overall. In addition, the number of subjects discontinued prior to receiving the 6<sup>th</sup> injection will be tabulated. The reasons for subjects discontinuing from radium-223 treatment, from study treatment, from active follow-up, and from survival follow-up will be tabulated by treatment group and overall. In addition, the number of subjects included in each analysis population will be displayed overall and by country and investigator.

#### 6.1.2 Demographics and Baseline Characteristics

All demographic and baseline characteristics will be summarized by cohort and overall for the phase 1b SAF population and by treatment group and overall for the phase 2 ITT population, using descriptive statistics such as frequency and proportion (for categorical variables), mean, median, standard deviation, minimum and maximum (for continuous variables).

These will include, but may not be limited to;

### Demographics

- Sex
- Age
- Age category (< 65 vs. ≥ 65 years)
- Race / Ethnic group
- Height (cm)
- Weight (kg)
- Body Mass Index, calculated as:

$$BMI = \frac{Weight (kg)}{[Height (m)]^2}$$

### Baseline disease characteristics

- Eastern Cooperative Oncology Group (ECOG) performance status
- Number of prior lines MM therapy (1 vs. >1)
- Prior Skeletal Related Events (SRE) (yes/no)
- Revised-ISS at initial diagnosis (Stage I, II and III)
- Revised-ISS at baseline (Stage I, II and III)
- Prior stem cell transplantation (yes, no)
- Serum albumin
- $\beta_2$  microglobulin
- Time since initial diagnosis to study entry (days)
- Time since first progression to study entry (days)
- Time since most recent progression to study entry (days)
- Time since first progression to most recent progression (days)
- Time since last bortezomib administration
- Baseline worst pain score (WPS) in the last 24 hours

### Baseline disease characteristics for phase 2 only

- Type of biological progression M-protein, PEP/SEP, FLC
- Bone marrow aspirate
- Response to prior therapy
- Time since last MM therapy

### 6.1.3 Medical history

Medical history will be summarized by MedDRA body system organ class (SOC) and preferred term (PT) for the Safety population by cohort and overall for the phase 1b part, and for the ITT population by treatment group and overall for the phase 2 part.

### 6.1.4 Drug Exposure and Compliance

Descriptive statistical summaries by cohort and overall for phase 1b and by treatment group and overall for phase 2 will be provided for the following variables using SAF:

- Number of radium-223 dichloride/placebo injections received. Summary statistics will include both descriptive statistics of number of injections received, and number and percent of subjects by number of injections received
- Number and percentage of radium-223 dichloride/placebo injection interruption or delays, reason for interruption or delays due to AEs, reasons other than AEs
- Number of interruption of bortezomib, reason for interruption or delays due to AEs, reasons other than AEs
- Dose reduction of bortezomib, reason for interruption or delays due to AEs, reasons other than AEs
- Cumulative dose or dose intensity of bortezomib
- Duration of study treatment for radium-223, bortezomib and dexamethasone, respectively: the date of the last dose of study treatment – date of the first dose of study treatment + 1
- Duration of treatment (overall): last dose date of last administered component – first dose date of first administer component +1.
- Total radioactivity of radium-223 received (sum of activity in all doses) in kBq and kBq/kg
- Planned dose, actual dose per day including interruptions/delays, actual dose per day excluding interruptions/delays, actual total amount of dose received, and percent of total planned dose for BOR and DEX will be summarized separately. Percent of total dose is calculated as actual daily dose per day including interruptions/delays, divided by the planned dose.

### 6.1.5 Prior and Concomitant Medications

All investigator-reported non-study medications taken before and/or during the study will be coded using the Bayer modified version of World Health Organization Drug Dictionary (WHO-DD) September 2017 and the Anatomical Therapeutic Chemical (ATC) classification system. Coding will include the drug class and generic drug name.

Non-study medications taken before, during and/or after the study will be categorized as prior medications, concomitant medications, and post treatment medications. They are defined based on when the medications are taken vs the first and last dosing dates of study treatment (radium-223 dichloride, BOR, or DEX). There are 8 scenarios see [Table 6–1](#).

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

All medications will be listed, including verbatim descriptions and coded terms, and flags for prior/concomitant/post treatment medications. Prior, concomitant, and post treatment medications will be summarized using frequencies of subjects reporting each drug category and generic drug name. For each subject, multiple records of the same concomitant medication will be counted once within a drug class and generic drug name.

The same approach will be used to summarize anti-cancer therapies (systemic anti-cancer therapy, radiotherapy as well as diagnostic and therapeutic procedures) in the Safety population.

**Table 6–1 Medication Classification**

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

C1= medication started before study drug administration and ended on or before study drug administration

C2= medication started before study drug administration and ended during study drug administration

C3= medication started on or after study drug administration and ended after study drug administration

C4= medication started before study drug administration and ended after study drug administration

C5= medication started on or after study drug administration and ended after study drug administration

C6= medication started on or after study drug administration and ended before or on the same date as end of study drug administration

Note: C7and C8 are covered under C6.

## 6.2 Efficacy

For phase 1b part, efficacy analyses will be performed for SAF. All efficacy analyses for phase 2 part will be performed for ITT population. Censoring for time to event endpoints (DOR, PFS, OS, time to SSE, time to pain progression, and time to opioid use) will be listed in each subsection accordingly, which follow the rules from the Xofigo project plan [3].

### 6.2.1 Primary Efficacy Analyses

#### Phase 1b

There is no planned primary efficacy analysis with phase 1b part.

#### Phase 2

Fisher’s exact test will be used for evaluating response rate difference between treatment groups. Response rate difference and its exact 95% CI will be provided. The descriptive

statistics will also be used to summarize response rates (combined CR+VGPR) , as determined by IMWG uniform response criteria. The number of responders, percentage of responders and 95% confidence interval [CIs]) will be presented by treatment arm. All ITT subjects with baseline disease assessment will be included as denominator in the response rate calculation.

All subjects who do not meet the criteria for a combined CR+VGPR by the analysis cutoff date will be considered non-responders.

## **6.2.2 Secondary Efficacy Analyses**

### **Phase 1b**

Anti-multiple myeloma response assessment (CR, VGPR, PR, etc.) will be summarized using descriptive statistics (i.e., number and percentage of subjects) by dose level cohort and overall. All subjects with baseline disease assessment will be included as denominator in the response rate calculation.

### **Phase 2**

Secondary efficacy analyses in phase 2 part comparing the effects of radium-223 dichloride dose group versus placebo in addition to background treatment with BOR plus DEX on the following set of endpoints will be performed: objective response rate (ORR) as determined by IMWG uniform response criteria, duration of response (DOR), progression-free survival (PFS), and overall survival (OS).

Fisher's exact test as described in Section 6.2.1 for primary efficacy endpoint will be used for evaluating ORR difference between treatment groups.

Time to event endpoints (DOR, PFS, OS) will be summarized by treatment arm using Kaplan-Meier estimates [1]. Kaplan-Meier curves will be generated, and median survival time together with the 25<sup>th</sup> and 75<sup>th</sup> percentiles and associated Brookmeyer-Crowley [2] 95% CIs will be presented. Log-rank test will be applied to test difference between treatment groups. The treatment effect (hazard ratio) will be estimated using the Cox proportional hazards regression model with treatment group as a factor.

#### **6.2.2.1 Objective response rate (ORR)**

The objective response rate is defined as the proportion of subjects in the analysis population who have complete response (CR), partial response (PR) and very good partial response [VGPR]) during the course of the study.

#### **6.2.2.2 Duration of Response (DOR)**

The duration of response (DOR) for a given subject, as determined by IMWG response criteria, will be defined as the time (in months) from the date of first response to treatment

(CR, VGPR or PR) to the date of disease progression or death. Duration of response is only applied to subjects with a disease response of at least PR.

The censoring rules for duration of response are summarized in Table 6-2.

**Table 6-2: Duration of Response censoring rules**

Situation (all situations occur after first response to treatment, i.e., CR, VGPR or PR)	End Date	Censored	Reason for Censoring
Subject had a disease assessment of PD (no more than one missed disease assessment)	Date of first PD	No*	N/A
Subject discontinued from study for other than PD or death	Last disease assessment without PD	Yes	Subject discontinued from study due to a reason other than PD or death
Death during the study (no more than one missed disease assessment) without disease PD assessment before death	Date of death	No*	N/A
Subject discontinued from study due to PD, but no documented PD date	Date of last disease assessment	Yes	Subjects discontinued from study due to PD, but no documented date of PD
Subject still on study at the time of data cutoff without PD	Last disease assessment before data cutoff	Yes	Subjects is still alive without PD
Death or PD after more than one missed or incomplete disease assessment	Date of last disease assessment before missed or incomplete assessment	Yes	Missed more than one disease assessment
New systemic anticancer treatment started prior to PD or death	Date of last disease assessment before starting new systemic anticancer treatment	Yes	New systemic anticancer treatment started

\*Earliest end dates in the above table are used in calculating the DOR.

### 6.2.2.3 Progression-free survival (PFS)

For a given subject, PFS will be defined as the time (in months) from the date of randomization to the date of disease progression or death, whichever occurs first.

The censoring rules for progression free survival is summarized in Table 6-3.



**Table 6-3: Progression Free Survival censoring rules**

<b>Situation</b>	<b>End Date</b>	<b>Censored</b>	<b>Reason for Censoring</b>
No baseline or post-baseline disease assessment	Randomization date	Yes	No baseline or post-baseline disease assessment
Subject had a disease assessment of PD (no more than one missed disease assessment)	Date of first PD	No*	N/A
Subject discontinued from study for other than PD or death	Last disease assessment without PD	Yes	Subject discontinued from study due to a reason other than PD or death
Death during the study (no more than one missed disease assessment) without disease PD assessment before death	Date of death	No*	N/A
Subject discontinued from study due to PD, but no documented PD date	Date of last disease assessment	Yes	Subjects discontinued from study due to PD, but no documented date of PD
Subject still on study at the time of data cutoff without PD	Last disease assessment before data cutoff	Yes	Subjects is still alive without PD
Death or PD after more than one missed disease assessment	Date of last disease assessment before missed assessment	Yes	Missed more than one disease assessment
New systemic anticancer treatment started prior to PD or death	Date of last disease assessment before starting new systemic anticancer treatment	Yes	New systemic anticancer treatment started

Progression Free Survival (PFS) = End Date – Date of Randomization +1

\*Earliest end dates in the above table are used in calculating the PFS.

#### 6.2.2.4 Overall Survival (OS)

Overall survival for a given subject will be defined as the time (in months) the date of randomization to the date of the subject's death.

The censoring rules for overall survival are summarized in Table 6-4.

**Table 6-4: Overall Survival censoring rules**

<b>Situation</b>	<b>End Date</b>	<b>Censored</b>	<b>Reason for Censoring</b>
Death on or prior to database cutoff date	Date of Death	No	N/A
Death occurred or subject known to be alive after database cutoff date	Database cut-off date	Yes	Alive as of data cut-off date
Last known alive date is on or prior to data cutoff date	Last known alive date (LKAD)	Yes	The reasons are based on the study CRF: e.g., subjects lost to follow up, withdrew consent etc.

The date of last known alive will be estimated based on the date that subject was reported alive from the post-treatment visits, date of disease progression assessment in the post-treatment visits, start/end date of post-study cancer therapy in the post-treatment visit, last available vital sign measurement, performance status assessment, or physical exam, CT scan, clinical visit for disease progression, whichever is the last.

### 6.2.3 Exploratory Efficacy Analysis

There is no planned exploratory efficacy analysis with phase 1b part. Exploratory analysis for time to event endpoints (time to SSE, time to pain progression, and time to opioid use) for phase 2 part will be summarized by treatment arm using Kaplan-Meier estimates. Kaplan-Meier curves will be generated, and median survival time together with the 25<sup>th</sup> and 75<sup>th</sup> percentiles and associated Brookmeyer-Crowley 95% CIs will be presented. Log-rank test will be applied to test difference between treatment groups.

#### 6.2.3.1 Time to first on-study SSE

The time to first on-study SSE is defined as the time (months) from the date of randomization to the date of the earliest occurrence of one of the following:

- the use of external beam radiotherapy (EBRT) to relieve skeletal symptoms
- the occurrence of new symptomatic pathological bone fractures (vertebral or non-vertebral)
- the occurrence of spinal cord compression
- a tumor-related orthopedic surgical intervention

The censoring rules for time to first on-study SSE is summarized in Table 6-5.

**Table 6-5: Time to first on-study SSE censoring rules**

<b>Situation</b>	<b>End Date</b>	<b>Censored</b>	<b>Reason for Censoring</b>
No post-baseline SSE assessment	Date of Randomization	Yes	No post-baseline SSE assessment
Subject had an SSE event	Date of first SSE	No	N/A
No SSE, and at least one on-study SSE assessment	Last SSE assessment	Yes	No SSE

### 6.2.3.2 Time to pain progression

Pain intensity is measured using the Brief Pain Index-Short Form (BPI-SF) at scheduled visits. In addition, 24-hour analgesic usage and opioid usage for each subject will be recorded on the eCRF. The time to pain progression is defined for evaluable subjects (i.e., subjects with a worst pain score of  $\leq 8$  in the BPI-SF at the baseline assessment) as the time (in days) from date of randomization until occurrence of the first post-baseline pain progression event.

Pain progression is defined as the occurrence of either a pain increase or an increase in pain management with respect to baseline, whichever occurs first.

A pain increase is defined as an increase of 2 or more points in the BPI-SF “worst pain in 24 hours” score from baseline observed at 2 consecutive evaluations  $\geq 28$  days apart.

An increase in pain management is defined as follows:

- For subjects not on opioids at baseline, initiation of short- or long-acting opioid use for pain will constitute an increase in pain management
- For subjects taking weak opioid at baseline, the initiation of any strong opioid or adding an additional weak opioid would be considered an increase in pain management.
- For subjects taking strong opioid at baseline, the initiation of an additional strong opioid (in addition to previous strong opioid) would be considered an increase in pain management.

The censoring rules for time to pain progression are summarized in Table 6-6.

**Table 6-6: Time to pain progression censoring rules**

<b>Situation</b>	<b>End Date</b>	<b>Censored</b>	<b>Reason for Censoring</b>
No post-baseline assessment	Date of Randomization	Yes	No post-baseline assessment
Did not experience pain progression by the cutoff date	Last post-baseline pain assessment date	Yes	Did not experience pain progression
Subject experienced a pain progression by the cutoff date	Date of first observation of pain progression	No	N/A

### 6.2.3.3 Change in MM disease markers

Change in PEP/SEP M-protein will be summarized for each visit and the final visit. Change in FLC will be summarized for each visit and the final visit for those subjects evaluated by FLC.

Changes in plasmocytes in bone marrow aspirates will be summarized by each visit and the final visit.

Descriptive statistics will be presented for the baseline and post-baseline of those markers at each visit. Mean change (95% CI) from baseline and mean difference between radium-223 dichloride dose group and placebo group (95% CI) will be examined based on ANCOVA model with treatment group as a factor, baseline M-protein as covariate.

### 6.2.3.4 Time to first opioid use for cancer pain

Time to opioid use for cancer pain is defined as the interval from the date of randomization to the date of opioid use. Subjects who have no opioid use at the time of analysis will be censored at the last known date of no opioid use. Subjects with no on-study assessment or no baseline assessment will be censored at the date of randomization.

The censoring rules for time to first opioid use for cancer pain are summarized in Table 6-7.

**Table 6-7: Time to first opioid use for cancer pain censoring rules**

<b>Situation</b>	<b>End Date</b>	<b>Censored</b>	<b>Reason for Censoring</b>
No on-study assessment or no baseline assessment	Date of Randomization	Yes	No on-study assessment or no baseline assessment
No opioid use at the time of datacut	Last known date of no opioid use	Yes	No opioid use
Used opioid on or prior to datacut	The date of first observation of opioid use	No	N/A

Analgesic use will be captured via three methods:

- Prior and concomitant medication case report form, where the physician records the analgesic medication prescribed to manage pain.
- 24 hour analgesic consumption case report form, in which all analgesic medication taken in the last 24 hours, including dose of medication and number of pills consumed, will be recorded. Patients will be asked to bring all of their medication to the clinic visit and complete this form with the assistance of the clinician.
- Opiate use case report form, in which the information of opiate pain medication since last assessment was collected.

For the time to first opioid use for cancer pain, the earliest date of the first opioid use recorded via any of the following three methods will be used: the 24 hour analgesic consumption case

report form, prior and concomitant medication case report form or opiate use case report form.

Patients who had opioid analgesics at baseline are not eligible for the analysis of this endpoint.

### **6.3 Pharmacokinetics / pharmacodynamics**

No PK measurement and analysis will be performed in this study.

### **6.4 Safety**

All subjects enrolled in phase 1b study will be evaluated for the occurrence of DLTs during the DLT observation window, which is defined as the time from the first dose of study treatment through 21 days after administration of the second dose of radium-223 dichloride. The definition of DLT can be referred to protocol Section 6.3.1. The number and percentage of DLTs will be summarized by cohort.

Safety variables will be analyzed using frequency tables and descriptive statistics by dose level cohort for phase 1b part and by treatment arm for phase 2 part. For phase 2 part, primary safety analysis will include treatment-emergent safety events as of database cutoff when all subjects completed end of radium-223 visit, or discontinued early. Final safety analysis will be performed after the last followed patient completes active follow-up 2 years following the last dose of radium-223 dichloride treatment (i.e., end of treatment visit). Final safety analysis will be reported in the Clinical Study report. In addition, safety events emerging during long-term follow-up will not be reported in the study Clinical Study Report, it will be reported in separate document(s).

All safety analyses will be performed for SAF.

#### **6.4.1 Adverse events**

All adverse events (AE) will be coded according to the latest Medical Dictionary for Regulatory Activities (MedDRA) Version as of database lock for the applicable analysis. The MedDRA version used will be documented in the CSR. The intensity of an AE will be documented using the NCI-CTCAE v4.03.

##### Adverse Events of Interest

A summary table and a listing will be provided for subjects who experienced following adverse events: hematological AEs (neutropenia, thrombocytopenia, etc), bone fractures, and secondary malignancies.

##### Pre-treatment AEs

A pre-treatment adverse event is defined as any event that started before the start of study drug administration.

### Treatment-emergent AEs

A treatment-emergent adverse event (TEAE) is defined as any event arising or worsening after start of study drug administration until the end of the treatment period. The treatment period is defined from the day of the first dose of radium-223 dichloride/placebo to 30 days after last dose of study treatment (radium-223 dichloride/placebo, BOR, or DEX, whichever is last)

### Post-treatment AEs

A post-treatment adverse event is defined as any event arising or worsening after the end of the treatment period.

An overall summary of AEs will be provided to present the number and percentage of subjects with

- any pre-treatment AEs, TEAEs, or post-treatment AEs
- any study treatment related TEAEs
- any serious TEAEs,
- any serious treatment related TEAEs
- any CTCAE Grade  $\geq 3$  TEAEs
- any TEAEs leading to drug modification
- any TEAEs leading to discontinuation
- any TEAEs leading to death
- any AEs leading to drug modification or discontinuation

TEAEs and post-treatment AEs will be summarized by MedDRA system organ class and preferred term. For each subject, multiple occurrences of the same event will be counted once within a system organ class and preferred term. For post-treatment AE tables, only treatment-related AEs, serious treatment-related AEs will be reported.

The same summaries will be repeated for related TEAEs, serious TEAEs, serious related TEAEs, CTCAE Grade  $\geq 3$  TEAEs, TEAEs leading to drug modification or discontinuation, TEAEs leading to death and any AEs leading to drug modification or discontinuation.

The maximum severity of the TEAEs and post-treatment AEs will be summarized according to the NCI-CTCAE toxicity criteria. For each subject, multiple occurrences of the same event will be counted once at their maximum severity within a system organ class and preferred term.

TEAE will also be summarized by NCI-CTCAE toxicity criteria and relationship to study medication.

Long-term follow-up (LTFU) AEs are AEs arising during the long-term follow-up period. LTFU AEs will not be included in safety analysis summary tables for this study, but will be included in listings.

Data listings will be produced for all AE recorded in the study. Verbatim descriptions and coded terms will be listed for all AEs.

Serious adverse events (SAEs), deaths, AEs leading to discontinuation and AEs of interest will each have a separate listing.

#### **6.4.2 Deaths**

Deaths reported during the study period will be summarized by timing and primary reason by cohort and overall for phase 1b part and by treatment group and overall for phase 2 part.

#### **6.4.3 Clinical Laboratory Data**

The following laboratory parameters will be summarized:

##### **Hematology:**

Hematocrit, Hemoglobin (Hb), platelet counts, RBC counts, white blood cell (WBC) counts, and WBC differential.

##### **Chemistry:**

Sodium, potassium, chloride, calcium (and albumin-adjusted calcium), ALT, AST, lactate dehydrogenase (LDH), total ALP, serum creatinine, blood urea nitrogen (BUN; or urea), bilirubin (total), alkaline phosphatase, glucose, phosphate, and albumin.

##### **Coagulation panel:**

PT, PTT, and INR of PT

Toxicity grading of laboratory values will be based on NCI CTCAE version 4.03. Descriptive statistics (number of observations, mean, standard deviation, minimum, median and maximum values) will be presented for quantitative clinical laboratory tests, their changes from baseline (including baseline value), and their percent changes from baseline at applicable visits. Graphical presentations will be generated for each laboratory parameter by visit to investigate trends over time and outliers in the data. In addition, changes from baseline will be summarized in shift tables according to baseline and worst CTCAE grades. Frequency tables will be provided for qualitative lab data.

If more than one assessment occurred at any visit (i.e. repeat samples taken), the last valid (non-missing) value will be used in the summaries. Unscheduled laboratory data will be listed but will not be included in the summary tables. However, all visits will be used in CTCAE summary tables.

#### 6.4.4 Other Safety Analyses

Vital sign measurements (systolic and diastolic blood pressure, heart rate, body temperature, respiration rate) and ECG will be summarized by visit. Assessments will be summarized with descriptive statistics at each applicable visit and for change from baseline to post-baseline assessments.

Body weight will be collected at screening and before each dosing. Body weight will be listed by visit.

Abnormal physical examination findings will be recorded either as medical history or as adverse events. There will be no separate listing for physical examination.

Pregnancy testing will be performed at the Screening and scheduled visits. All the pregnancy reported during the study will be included in the SAE listing, no separate listing will be generated for pregnancy.

The new primary malignancies will be summarized by cancer type (n, percent), and reported through listings.

#### 6.5 Maximum tolerated dose (MTD)/Recommended phase 2 dose (RP2D)

MTD/RP2D will be evaluated by the sponsor in consultation with the investigators during the conduct of the dose escalation part of the study. MTD/RP2D will be determined using the incidence of DLTs during the DLT observation window, which is defined as the time from the first dose of study treatment through 21 days after administration of the second dose of radium-223 dichloride. The decision about MTD/RP2D will be made after all subjects in last dose escalation cohort are treated with at least 2 doses for radium-223 dichloride and observed for 21 days after second dose. If no DLT is observed in Phase 1b part, a RP2D will be selected based on safety and efficacy data obtained during the phase 1b part.

Individual listings of DLTs will be presented by dose level cohort for phase 1 part with treatment cycle of onset, serious or nonserious, CTCAE v4.03 grade and term, and MedDRA terms provided for each DLT. Dose Evaluable Set (DES) will be used for the analyses.

#### 6.6 Analysis of Other Endpoints

All the analyses listed in this section will be performed by cohort for phase 1b part for SAF and by treatment group for phase 2 part for ITT population.

##### 6.6.1 ECOG Performance Status

The ECOG-PS is to be assessed every Day 1 of BOR before injection. The number and percentage of subjects in each ECOG-PS category will be presented by visit. Changes from



baseline in ECOG-PS score will be summarized in shift tables according to baseline and post-baseline assessment by visit.

### **6.6.2 Brief Pain Index – Short Form (BPI-SF)**

BPI-SF is composed of 9 questions asking whether subject experiences pain, location and severity of the pain. Subject will be asked to rate the pain they experienced with 0 as no pain and 10 for as pain as imagine. Score for Question 3 “worst pain in the last 24 hours” will be used for calculating time to pain progression as described in Section 6.2.3. A data listing will be presented for BPI-SF assessments at each scheduled visit.

### **6.6.3 Resource Utilization Questionnaire**

Information on healthcare resource use that is associated with the management of AEs as well as subject monitoring will be collected by resource utilization questionnaire. A data listing will be presented for RUQ at each scheduled visit.

### **6.6.4 Long Term Safety**

Safety data arising during the long-term safety follow-up period in this study will be presented through listings. In the event subjects in this study transfer into a long term safety follow-up study for all or a portion of their long-term follow-up period, any data not collected in this study is outside the scope of this SAP.

## **6.7 Examination of Subgroups**

There will be no planned subgroup analysis for this study.

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

Changes in the planned statistical analysis from protocol:

- 1) “changes in bone lesions” were removed from the analysis due to no data will be collected.
- 2) Analysis for bone biomarkers, ctDNA, CTCs, and MM related cytokines were removed as they will be described in a separate SAP.
- 3) Time to first on-study SSE, time to pain progression, and time to first opioid use are defined as the time from the date of randomization, instead of from the first dose of radium-223 dichloride.



## 8. References

1. Kaplan E and Meier P. Nonparametric estimation from incomplete observations. *J Amer Statist Assn* 53:457–481 (1958)
2. Brookmeyer R and Crowley J. A confidence interval for the median survival time. *Biometrics* 38:29-41 (1982)
3. Xofigo Biostatistical project Plan version 2.0 dated 12 December 2016.
4. \_Xofigo Biostatistical Project Plan\_v2\_07Dec2016 final\_addendum.docx dated 17 May 2018.