Document Type:	Statistical Analysis Plan
Official Title:	A phase II randomized, double-blind, placebo-controlled trial of radium-223 dichloride versus placebo when administered to metastatic HER2 negative hormone receptor positive breast cancer subjects with bone metastases treated with hormonal treatment background therapy
NCT Number:	NCT02258464
Document Date:	12 Aug 2019

Statistical Analysis Plan (Amendment/Supplement) Approval Form

Study Number (Bay No./IMP no.)*	BAY 88-8223 / 16298
Statistical Analysis Plan (SAP) Version and Date	Version 4.0, 12 AUG 2019

* if no IMPACT number is available, refer to the approved Study Concept

I have read and approve the SAP/SAP Amendment referred above.

	Name	Signature	Date
Author:			
Study Statistician *<plus affiliation, if not Bayer>*	PPD	PPD	12 Aug 2019
Approved by:			
Project Statistician/Clinical Sciences Statistician[1)]	PPD	PPD	20 Aug 19
Senior Clinical Development Leader[2) 4)]/Early Clinical Leader/ Clinical Pharmacology Leader[3)]	PPD		
Clinical Development Leader[2) 4)]/ Early Clinical Leader/ Clinical Pharmacology Leader[3)]	PPD	PPD	12-Aug-19
Study Statistical Analyst *<plus affiliation, if not Bayer>*	PPD		
Medical Writer *<plus affiliation, if not Bayer>*	PPD		
Global Health Economics & Outcomes Research Project Leader	PPD		

1) CS STAT for CS owned studies
2) the CDL may delegate his review and approval task
3) ECL for Phase 1/2a studies owned by EM and CPL for Phase 1/2a studies owned by CS
4) MAR for studies managed by MA

Statistical Analysis Plan (Amendment/Supplement) Approval Form

Study Number (Bay No./IMP no.)*	BAY 88-8223 / 16298
Statistical Analysis Plan (SAP) Version and Date	Version 4.0, 12 AUG 2019

* if no IMPACT number is available, refer to the approved Study Concept

I have read and approve the SAP/SAP Amendment referred above.

	Name	Signature	Date
Author:			
Study Statistician *<plus affiliation, if not Bayer>*	PPD		
Approved by:			
Project Statistician/Clinical Sciences Statistician[1)]	PPD		
Senior Clinical Development Leader[2) 4)]/Early Clinical Leader/ Clinical Pharmacology Leader[3)]	PPD	PPD	13 Aug 2019
Clinical Development Leader[2) 4)]/ Early Clinical Leader/ Clinical Pharmacology Leader[3)]	PPD		
Study Statistical Analyst *<plus affiliation, if not Bayer>*	PPD		
Medical Writer *<plus affiliation, if not Bayer>*	PPD		
Global Health Economics & Outcomes Research Project Leader	PPD		

1) CS STAT for CS owned studies
2) the CDL may delegate his review and approval task
3) ECL for Phase 1/2a studies owned by EM and CPL for Phase 1/2a studies owned by CS
4) MAR for studies managed by MA

Statistical Analysis Plan (Amendment/Supplement) Approval Form

Study Number (Bay No./IMP no.)*	BAY 88-8223 / 17096
Statistical Analysis Plan (SAP) Version and Date	Version 4.0, 12 AUG 2019

* if no IMPACT number is available, refer to the approved Study Concept

I have read and approve the SAP/SAP Amendment referred above.

	Name	Signature	Date
Author:			
Study Statistician <*plus affiliation, if not Bayer*>	PPD		
Approved by:			
Project Statistician/Clinical Sciences Statistician[1)]	PPD		
Senior Clinical Development Leader[2) 4)]/Early Clinical Leader/ Clinical Pharmacology Leader[3)]	PPD		
Clinical Development Leader[2) 4)]/ Early Clinical Leader/ Clinical Pharmacology Leader[3)]	PPD		
Study Statistical Analyst <*plus affiliation, if not Bayer*>	PPD	PPD	8/23/2019
Medical Writer <*plus affiliation, if not Bayer*>	PPD		
Global Health Economics & Outcomes Research Project Leader	PPD		

1) CS STAT for CS owned studies
2) the CDL may delegate his review and approval task
3) ECL for Phase 1/2a studies owned by EM and CPL for Phase 1/2a studies owned by CS
4) MAR for studies managed by MA

Statistical Analysis Plan (Amendment/Supplement) Approval Form

Study Number (Bay No./IMP no.)*	BAY 88-8223 / 16298
Statistical Analysis Plan (SAP) Version and Date	Version 4.0, 12 AUG 2019

* if no IMPACT number is available, refer to the approved Study Concept

I have read and approve the SAP/SAP Amendment referred above.

	Name	Signature	Date
Author:			
Study Statistician *<plus affiliation, if not Bayer>*	PPD		
Approved by:			
Project Statistician/Clinical Sciences Statistician[1)]	PPD		
Senior Clinical Development Leader[2) 4)]/Early Clinical Leader/ Clinical Pharmacology Leader[3)]	PPD		
Clinical Development Leader[2) 4)]/ Early Clinical Leader/ Clinical Pharmacology Leader[3)]	PPD		
Study Statistical Analyst *<plus affiliation, if not Bayer>*	PPD		
Medical Writer *Covance*	PPD	PPD	14 Aug 2019
Global Health Economics & Outcomes Research Project Leader	PPD		

1) CS STAT for CS owned studies
2) the CDL may delegate his review and approval task
3) ECL for Phase 1/2a studies owned by EM and CPL for Phase 1/2a studies owned by CS
4) MAR for studies managed by MA

Statistical Analysis Plan (Amendment/Supplement) Approval Form

Study Number (Bay No./IMP no.)*	BAY 88-8223 / 16298
Statistical Analysis Plan (SAP) Version and Date	Version 4.0, 12 AUG 2019

* if no IMPACT number is available, refer to the approved Study Concept

I have read and approve the SAP/SAP Amendment referred above.

	Name	Signature	Date
Author:			
Study Statistician *<plus affiliation, if not Bayer>*	PPD		
Approved by:			
Project Statistician/Clinical Sciences Statistician[1)]	PPD		
Senior Clinical Development Leader[2) 4)]/Early Clinical Leader/ Clinical Pharmacology Leader[3)]	PPD		
Clinical Development Leader[2) 4)]/ Early Clinical Leader/ Clinical Pharmacology Leader[3)]	PPD		
Study Statistical Analyst *<plus affiliation, if not Bayer>*	PPD		
Medical Writer *<plus affiliation, if not Bayer>*	PPD		
Global Health Economics & Outcomes Research Project Leader	PPD	PPD	12 AUG 2019

1) CS STAT for CS owned studies
2) the CDL may delegate his review and approval task
3) ECL for Phase 1/2a studies owned by EM and CPL for Phase 1/2a studies owned by CS
4) MAR for studies managed by MA

Integrated Statistical Analysis Plan

A phase II randomized, double-blind, placebo-controlled trial of radium-223 dichloride versus placebo when administered to metastatic HER2 negative hormone receptor positive breast cancer subjects with bone metastases treated with hormonal treatment background therapy

For this study, the SAP and subsequent SAP amendment were released as follows:

- Original statistical analysis plan, Version 1.0, dated 29 APR 2016
- Amendment 1 (based on integrated protocol amendment7) forming integrated statistical analysis plan Version 2.0, dated 25 OCT 2017
- Amendment 2 (based on integrated protocol amendment 8) forming integrated statistical analysis plan version 3.0, dated 4 JAN 2019
- Amendment 3 (based on integrated protocol amendment 8) forming integrated statistical analysis plan version 4.0, dated 12 AUG 2019

This document integrates the original statistical analysis plan and all global amendments.

A phase II randomized, double-blind, placebo-controlled trial of radium-223 dichloride versus placebo when administered to metastatic HER2 negative hormone receptor positive breast cancer subjects with bone metastases treated with hormonal treatment background therapy

Short title: Study of radium-223 dichloride versus placebo and hormonal treatment as background therapy in subjects with bone predominant HER2 negative hormone receptor positive metastatic breast cancer

BSP Study drug	BAY 88-8223 / Radium-223 dichloride / Xofigo®
Study purpose:	This is a randomized, double-blind, placebo-controlled phase II study to evaluate the clinical benefit of radium-223 dichloride versus placebo when administered to metastatic HER2 negative hormone receptor positive breast cancer subjects with bone metastases treated with hormonal treatment background therapy

Clinical study phase:	II	**Date:**	12AUG2019
Study No.:	BAY 88-8223 / 16298	**Version:**	V4.0

Author: PPD
PPD
PPD

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.

The approval of the Statistical Analysis Plan is documented in a separate Signature Document.

Table of Contents

Table of tables

Abbreviations

AE	Adverse event
ALP	Alkaline phosphatase
BHA	Denosumab or bisphosphonates
BMI	Body Mass Index
BPI-SF	Brief Pain Inventory – Short Form
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events; version 4.03
EBRT	External beam radiation therapy
ECOG	Eastern Co-operative Oncology Group
ECOG PS	Eastern Co-operative Oncology Group Performance Status
eCRF	Electronic case report form
EOT	End of Treatment
GCP	Good Clinical Practice
HR	Hazard ratio
ICF	Informed consent form
IDMC	Independent data monitoring committee
ITT	Intent-to-treat
IV	Intravenous
IXRS	Interactive Voice/Web Response System
kBq	KiloBecquerel; SI unit of radioactivity
kg	Kilogram
mCi	Millicuries
MedDRA	Medical Dictionary for Regulatory Activities
mL	Milliliter
MRI	Magnetic resonance imaging
NCI	National Cancer Institute
OS	Overall survival
PD	Progressive disease
PS	Performance status
rPFS	Radiological progression-free survival
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	Serious adverse event
SAP	Statistical analysis plan
SAS	Statistical analysis software
SSE	Symptomatic skeletal event
SSE-FS	Symptomatic skeletal event-free survival
WHO	World Health Organization
WHO-DD	World Health Organization Drug Dictionary
WPS	Worst pain subscale

Definitions of terms

Radium-223 dichloride	The investigational product, a targeted alpha particle emitting radiopharmaceutical, is a ready-to-use solution for intravenous injection containing the drug substance radium-223 dichloride. The active moiety is the alpha particle emitting nuclide radium-223, present as a divalent cation ($^{223}Ra^{2+}$).
Dose	Doses are given as kiloBecquerel (kBq) per kilogram body weight, with the corresponding dose given in millicurie (mCi) per kilogram in parenthesis. The term “dose” is used to describe the quantity of radioactivity from radium-223 administered.

1. Introduction

Presence of estrogen receptor (ER) and/or progesterone receptor (PR) is one of the most important prognostic/predictive factors in breast cancers. Patients with hormone receptor positive disease and HER2-negative disease are candidates for endocrine therapy. Bone is a frequent site of metastatic spread with approximately 65% to 75% of patients with metastatic breast cancer having skeletal involvement [1]. Due to the disrupted bone remodeling process, patients with metastatic bone lesions are at risk of increased morbidity including skeletal-related events (SRE), such as bone pain requiring intervention (i.e., radiotherapy or surgery), pathologic fractures, spinal cord compression, as well as symptomatic hypercalcemia and bone marrow infiltration. These events will ultimately impair the patient's quality of life (QoL) and functional independence.

Radium-223 dichloride solution for injection is a novel alpha particle-emitting radiopharmaceutical. The bone targeting property of radium-223 is similar to that of other earth alkaline elements, like calcium or strontium-89. However, the radiation characteristics of an alpha particle-emitting radionuclide appear to be more advantageous than of a beta-emitting radionuclide. The alpha particle range from radium-223 dichloride is less than 100 micrometers (less than 10 cell diameters), which limits damage to the surrounding normal tissue.

In the phase III, double-blind, randomized, BC1-06, ALSYMPCA (**Al**pharadin in **Sym**ptomatic **P**rostate **Ca**ncer) study, a total of 922 subjects with CRPC and symptomatic bone metastases were randomized to receive 6 injections of radium-223 dichloride (50 kBq/kg IV, based on NIST 2010 standardization [2]) or matching placebo every 4 weeks. Based on data of an interim analysis (n=809), radium-223 dichloride significantly improved overall survival (OS compared to placebo (the median OS was 14.0 versus 11.2 months, respectively; Hazard ratio [HR]= 0.695; p=0.00185). Symptomatic skeletal events (SSE) were lower in the radium-223 dichloride arm, and time to first SSE was significantly delayed (the median time to SSE was 13.6 months, versus 8.4 months, respectively; HR=0.610; p=0.00046). Adverse events of any grade were reported in 88% of the subjects who received radium-223 dichloride versus 94% in the placebo arm (Grade 3/4 AEs were reported for 51% and 59% of subjects, respectively).

In an open-label, multicenter, single-arm, phase IIa study (BC1-09), 23 subjects with metastatic breast cancer with bone dominant disease were administered 4 injections of radium-223 dichloride (50 kBq/kg IV, based on NIST 2010 standardization[2]) every 4 weeks. The primary efficacy endpoints were the change in urine levels of N-terminal telopeptide (NTX) and bone alkaline phosphatase (ALP) from baseline at Week 16. Median urine NTX levels were reduced by 20% (from 36 to 29 nmo1 bone collagen equivalents [BCE]/mmo1 creatinine; p=0.03) and 33% (from 36 to 23 nmo1 BCE/mmo1 creatinine; p=0.0124) at Week 8 and Week 16, respectively; 17/23 and 9/13 subjects (for whom data were available) had a decrease in urine NTX at Week 8 and Week 16, respectively. Median bone ALP levels were reduced by 33% (from 22.1 to 12.1 ng/mL; p=0.0001) at Week 8 and 42% (from 22.1 to 10.94 ng/mL; p=0.04) at Week 16. Bone ALP levels were reduced in 20/22 subjects at Week 8 and in 10/12 subjects (for whom data were available) at Week 16. Radium-223 dichloride was found to be safe and well tolerated.

The treatment options for patients with bone dominant metastasis of breast cancer are still limited.

Radium-223 dichloride has shown significant anti-tumor activity in a phase III study in subjects with bone predominant metastatic CRPC and in a phase II metastatic breast cancer study. The safety profile and tolerability for radium-223 dichloride appear to be acceptable in this study population.

For a more detailed introduction, please refer to the clinical study protocol.

The original protocol, version 1.0, is dated 13 MAY 2014. This statistical analysis plan (SAP) version 4.0 (amendment 3), describes the analyses to be included in the clinical study report for study 16298. It is based on the SAP version 1.0, dated 29 APR 2016, the version 2.0 (amendment 1), dated 25 OCT 2017, the version 3.0 (amendment 2), dated 04 JAN 2019 and the integrated clinical study protocol version 8.0 (amendment 8), dated 03 APR 2018.[1]

2. Study Objectives

The objective of this study is to assess the efficacy and safety of radium-223 dichloride in subjects with HER2 negative, hormone receptor positive breast cancer with bone metastases treated with hormonal treatment background therapy.

The primary endpoint is:

- symptomatic skeletal event-free survival (SSE-FS).

The secondary endpoints are:

- OS;
- time to opiate use for cancer pain;
- time to pain progression;
- pain improvement rate;
- time to cytotoxic chemotherapy;
- radiological progression free survival (rPFS);
- Safety, acute and long term, including new primary malignancies and hematopoietic reserve for tolerability of subsequent chemotherapy.

The exploratory endpoints are:

[1] As of Amendment 3.

- time to first on-study SSE;
- time to bone ALP progression;
- bone ALP response at Week 12 and end of treatment (EOT);
- bone-specific rPFS;
- resource utilization;
- biomarker assessments,
- time to visceral metastases onset.

3. Study Design

This study is an international phase II, randomized, double-blind, placebo-controlled, parallel group study of radium-223 dichloride versus placebo administered on top of the hormonal treatment and supportive care as background treatment in subjects with HER2 negative, hormone receptor positive breast cancer with bone metastases. Randomization will be stratified by:

- Geographical regions (Europe/North America [including Israel] versus Asia)
- Previous lines of hormone therapy in metastatic setting (1 versus 2 or more): for the purpose of counting the number of prior lines of hormone therapy, only a change of the hormone agent due to progression is counted as a new line of therapy. A switch of hormone therapy from one agent to another due to toxicity or other reasons (e.g., subject's preference) in absence of progressive disease (PD) at the time of switch will be counted as one line, although 2 different agents have been administered.
- Prior SREs (1 versus 2): for the purpose of prior SREs stratification, separate SREs are those that occur at least 21 days apart of each other. Any procedure which is related to an SRE, such as orthopedic surgery to treat a pathological bone fracture or multiple doses of radiation during a course of treatment, should not be counted as a separate event. In case of bone pain that occurs in several anatomical locations and requires separate external beam radiotherapy (EBRT) sessions, it should be counted as one event if the EBRT sessions are administered within a period of 21 days.

This study will be conducted at approximately 160 investigative study sites and approximately 227 subjects will be enrolled.

The study will comprise 4 periods: screening, randomization, treatment, and the follow-up period (active follow-up with clinic visits and active follow-up without clinic visits).

Investigational treatment consists of up to 6 cycles of radium-223 dichloride 55 kBq/kg body weight (based on updated 2015 NIST standardization[3][1]) (Arm A) or placebo (Arm B) each separated by an interval of 4 weeks. All subjects will receive hormonal treatment with a single agent and supportive care as background treatment according to the local standard of practice. These treatments will also continue after completion of radium-223 dichloride or placebo.

Subjects will be assessed for efficacy and safety endpoints at each treatment visit and will be evaluated every 12 weeks for radiological progression. Symptomatic skeletal events should be recorded until end of active follow-up, independent of whether the subject starts a new anti-cancer therapy (i.e., chemotherapy, other).

Subjects who discontinued or completed the radium-223 dichloride or placebo treatment, did not experience an SSE **and** radiological progression during the treatment period, had an EOT visit, and can travel to the clinic will enter an active follow-up period with clinic visits. During this period, clinic visits are to occur as follows:

- For subjects who have not experienced an SSE during the treatment period, visits will continue with the same frequency as during treatment (every 4 weeks ±7 days) until the subject has an SSE. After the occurrence of the SSE, the subjects will be switched to a frequency of visits every 12 weeks ±7 days at the next scheduled visit.

- For subjects who experienced an SSE during the treatment or this period, visits will occur every 12 weeks ±7 days. The subjects will continue to be followed for radiological progression and long-term safety.

- Subjects who miss 2 consecutive follow-up visits will be considered unable to travel to the site and will enter the active follow-up without clinic visits.

Subjects from the treatment period or the active follow-up period with clinic visits who can no longer travel to the clinic site or those who experienced an SSE and radiological progression will be followed for survival, treatment-related AEs and SAEs, and the initiation of other anti-cancer therapies, and SSEs, with phone calls as follows:

- For subjects who did not experience an SSE during treatment or the previous follow up period, the frequency of the phone calls will be every 4 weeks ±7 days until the subject has an SSE. After the occurrence of the SSE, the subjects will be switched to a frequency of phone calls every 12 weeks ±7 days at the next scheduled phone call.

- For subjects who experienced an SSE during treatment or the previous follow-up period, the frequency of the phone calls will be every 12 weeks ±7 days.

A schematic of the study design is presented in protocol Section 5.

[1] Updated dose per revised NIST standardization. The amount of radioactivity in the administered dose is unchanged from the 2010 standardization; only the numerical value of the dose changed due to the revised standard.

4. General Statistical Considerations

4.1 General Principles

The statistical evaluation will be performed by using the software package SAS version 9.2 or higher (SAS Institute Inc., Cary, NC, USA). Unless otherwise noted, data will be analyzed by descriptive statistical methods: The number of data available and missing data, mean, standard deviation, minimum, quartiles, median, and maximum will be calculated for metric data. Frequency tables will be generated for categorical data.

Definition of efficacy and safety endpoints, analysis strategies, structure of analysis datasets and layout of analysis data displays are following Bayer Healthcare Pharmaceuticals (BHP) standards: Xofigo Project Standards, the Therapeutic Area Oncology Standards (TAS) and the Global Medical Standards (GMS), respectively. Where the given ordering reflects the priority of the different standards, means specifications of the latter ones have to be followed only if not specified in standards mentioned before.

4.2 Handling of Dropouts

A "dropout" is defined as a subject who has been randomized and discontinues study participation prematurely for any reason. Subjects withdrawn from study treatment will not be replaced. Refer to Section 6.3 in the study protocol for withdrawal of subjects from study.

All efficacy analyses are based on the ITT population, which comprises all randomized subjects, including subjects who withdraw regardless of the reason for withdrawal. See following chapters for more details on deriving efficacy endpoints in case of missing data.

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 eCRF. Except as noted, missing data will not be imputed or carried forward in any statistical analysis.

Safety baseline is defined as the available value at last visit prior to or on first administration date of radium-223 dichloride. If values are missing at the baseline (Visit 2), data recorded at screening visit will be considered as baseline value. If screening record is also missing, the baseline value will be left as missing.

No imputation will be performed for missing lesion assessment and tumor response. For example, if a subject misses a scan visit and PD is documented at the next available scan visit, the actual visit date of the first documented PD will be used to calculate rPFS.

The partial missing date used for derivation of time to event, adverse event and concomitant medication partial missing start/stop dates will be imputed and the imputation rule will be specified in the analysis dataset specification.

4.4 Interim Analyses and Data Monitoring

4.4.1 Interim Analyses

An administrative interim data review will be performed when approximately 40 rPFS events are reached. The interim data review will be primarily focused on the rPFS, and the results will inform on future radium-223 clinical development plans in this indication. An independent unblinded data review committee will be formed.

No formal statistical testing will be performed for either SSE-FS or rPFS at the time of this interim review. Only summary statistics will be produced. The number of patients randomized, number censored, number of events, median rPFS and its 80% and 95% CI, and rPFS event rate at 6, 9 and 12 months will be presented. A Kaplan-Meier curve will be generated for each treatment group.

There is no plan to stop the trial due to superior efficacy; therefore, no alpha adjustment is applied for this interim look at the primary endpoint of SSE-FS. The study will remain blinded after the interim data review.

Treatment-related AEs, study drug-related TEAEs, TE, AEs of interest, and serious AEs, lab toxicities are planned to be reported. The summaries of the safety data will be completed for the Safety analysis population (SAF). [1]

4.4.2 Data Monitoring

An Independent Data Monitoring Committee (IDMC) will not be applied to this study.

4.4.3 Data Rules

A data cut-off date for the final analysis of SSE-FS will be based on the number of events required for the primary endpoint SSE-FS that have occurred.

4.5 Validity Review

The results of validity review meetings will be documented in the Validity Review Reports.

5. Analysis Sets

Intent-to-treat analysis (ITT) population

[1] As of Amendment 1.

The primary population for all efficacy analyses is the Intent-to-treat analysis (ITT) population, which is defined as all randomized subjects. Subjects will be analyzed as randomized.

Safety analysis (SAF) population

The population for safety analysis (SAF) population is comprised of all randomized subjects who received at least one dose of study medication (Radium-223 dichloride or placebo) and will be used for all safety analyses. Subjects will be analyzed as treated. Using a conservative approach, subjects randomized to the placebo arm will be analyzed under the Radium-223 dichloride arm if the subject received at least one dose of Radium-223 dichloride treatment whereas subjects randomized to the Radium-223 dichloride arm will still be analyzed under the Radium-223 dichloride arm regardless of receiving at least one dose of placebo. Only in a case where a subject randomized to the Radium-223 dichloride arm received only placebo treatment would the subject be analyzed under the placebo arm.

6. Statistical Methodology

The formal statistical analyses will be both descriptive and inferential. Summaries will be provided for each of the treatment groups.

6.1 Population characteristics

For this analysis, only descriptive statistics will be provided (no testing performed). The results will be displayed by treatment group for the ITT population.

6.1.1 Disposition of subjects

The number of subjects enrolled and included in each of the ITT and safety populations will be tabulated by region, country, and center. A summary table will also be presented for the number of subjects enrolled and the number and percentage of subjects in each of the defined populations. The reasons for subjects excluded from each of the analysis populations will also be tabulated. In addition, the number of subjects who were screened, randomized, treated and discontinued will be summarized by treatment group. Reasons for discontinuation of study treatment will be tabulated.

6.1.2 Demographic and baseline characteristics

All demographic and baseline characteristics will be summarized for the ITT population using descriptive statistics such as frequency and proportion (for categorical variables), mean, median and standard deviation (for continuous variables).

These will include, but may not be limited to:

- Demographics and baseline characteristics:
  Sex, race, ethnicity, age calculated at the date of randomization using date of birth, age

categories (≥55, <55) [1], Geographical region (Europe/North America [including Israel] versus Asia), baseline weight, height, body mass index, systolic blood pressure (BP), diastolic BP, heart rate, respiration rate.

- Baseline cancer characteristics:
  Eastern Cooperative Oncology Group (ECOG) performance status, menopausal status, histology, stage at initial diagnosis (TNM), grading (American Joint Committee on Cancer [AJCC]) at initial diagnosis, status of primary tumor at study entry, TNM classification of breast cancer at study entry, progesterone receptor status, estrogen receptor status, HER2/neu status (Immunohistochemistry[IHC]), HER2/neu status (Fluorescence in situ hybridization[FISH])/(Chromogenic in situ hybridization[CISH]) /(HER2 gene amplification ISH (other validated assay)) [2], time since initial diagnosis, time since initial diagnosis to metastatic disease, time since initial diagnosis to bone metastases, time since first progression, time since most recent progression, previous lines of hormone therapy in metastatic setting (1 versus 2 or more), and prior SREs (1 versus 2), cancer pain assessment[3].

6.1.3 Medical history

The dictionary for coding is MedDRA version 17 or most recent version. By treatment-group summary statistics (frequency and percentage) will be provided by system organ class and preferred term for ITT analysis set.

6.1.4 Prior, concomitant and post-treatment medications

The dictionary used for coding medications is World Health Organization Drug Dictionary (WHO-DD). The following categories of medications will be summarized

- Systemic anti-cancer therapies: frequency of subjects for each drug category, prior and post study medication.
- Diagnostic and Therapeutic Procedures: frequency of subjects by procedure, prior, during and post study medication.
- Radiotherapies: frequency of subjects by field and intent, prior, during therapy and post study medication.
- Myelosupressive systemic anti-cancer therapy during follow up: frequency of subjects for each drug category.
- Analgesics: frequency of subjects for each drug category, prior and post study medication.
- Denosumab or bisphosphonates (BHA): frequency of subjects for each drug category, prior study medication.

[1] As of Amendment 1.
[2] As of Amendment 1.
[3] As of Amendment 1.

- Other prior and concomitant medications: frequency of subjects for each drug category.

6.2 Efficacy Analysis

All comparisons of the treatment groups with respect to the primary and secondary efficacy variables will be based on the ITT population.

The duration for the time-to-event variables is calculated by event/censoring date – reference date + 1. Time-to-event analysis reference date is randomization date except for time to opiate use for cancer pain, time to pain progression, pain improvement rate. These pain efficacy endpoints baseline is defined as last available value prior to first administration of radium-223 dichloride since the data are not collected at screening visit. For all other efficacy endpoints, the baseline is defined as the last available value prior to or on the date of randomization.

6.2.1 Primary efficacy variables

6.2.1.1 Definition of SSE-FS:

The primary efficacy endpoint is symptomatic skeletal event-free survival (SSE-FS). It is defined as the time from date of randomization to occurrence of one of the following, whichever happens earlier:

(1) An on-study SSE, which is defined as:

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

(2) Death from any cause

Note: Any EBRT or orthopedic surgery related to a prior skeletal related event but administered after signature of the informed consent form (ICF) will not be counted as an on-study SSE.

For event date, the actual procedure date instead of date of symptom will be used to calculate the SSE-FS.

The censoring rules for SSE-FS are summarized in Table 1.

Table 1: Symptomatic Skeletal Event Free Survival (SSE-FS) censoring rules.

Situation	End Date	Censored	Reason for Censoring
No post-baseline SSE assessment and no death	Date of Randomization	Yes	No post-baseline SSE assessment and no death
Subject had an SSE event	Date of first SSE	No*	N/A
Death without prior SSE (<9 weeks between last SSE assessment and death) #	Date of death	No*	N/A
Death without prior SSE (≥ 9 weeks between last SSE assessment and death) #	Last SSE assessment before the missing SSE assessments#	Yes	≥ 9 weeks between last SSE assessments immediately prior to death
Neither SSE nor death at data cutoff	Last SSE assessment	Yes	Neither SSE nor death

Symptomatic Skeletal Event Free Survival (SSE-FS) = End Date – Date of Randomization +1
*The earliest end date in the table is used in calculating the SSE-FS
#: use randomization date instead of last SSE assessment date if no post-baseline SSE assessment.
SSE events immediately after missing SSE assessments are still counted as events in the analysis of SSE-FS.

6.2.1.2 Analysis of SSE-FS:

The primary analysis of SSE-FS will be performed when approximately 119 subjects have SSE-FS.

Since differences are expected between the values of the stratification variables entered by the investigator at the time of randomization (IVRS) and those derived from information collected on the eCRF, the analysis will be performed using both assignments to the strata. However, the primary stratified analyses for the efficacy endpoints will be based on the stratification information entered in IVRS. Stratification factors collected on the eCRF (considered the 'true' information) will be used for sensitivity analyses. In rare circumstances, if eCRF data cannot be accurately documented, then IVRS data will be used.

The null hypothesis that both treatment groups have the same SSE-FS distribution will be tested against the alternative hypothesis that the distribution of SSE-FS time in radium-223 dichloride is different from the placebo group.

H_0: $SSE\text{-}FS_{Xofigo} = SSE\text{-}FS_{Placebo}$, versus

H_A: $SSE\text{-}FS_{Xofigo} > SSE\text{-}FS_{Placebo}$[1]

[1] As of Amendment 1.

The SSE-FS will be compared using a stratified log-rank test with a one-sided alpha of 0.1 (2-sided alpha of 0.2), stratified by the same stratification factors as randomization: geographical regions, previous lines of hormone therapy in metastatic setting and prior SREs, based on the stratification data in IVRS.

The product-limit estimates of the survival distribution functions (Kaplan-Meier) will be presented for each treatment group: N, total censored, total failed, time to event (median and its 80% CI, 25th and 75th percentile, range), probability of event-free survial at pre-specified months. A Kaplan-Meier curve will be generated for each treatment group.

The hazard ratio (radium-223 dichloride / placebo) for SSE-FS and its 80% and 95% confidence intervals (CIs) will be calculated using a univariate Cox model, stratified by the same factors as stated above.

The Kaplan-Meier (KM) estimates for SSE-FS and KM curves will also be presented for each treatment group. The median and SSE-FS rates at time points such as 3 months, 6 months, etc., together with corresponding 80% CIs will also be calculated by treatment group. Subgroup analysis will be performed for subjects who have no opiate use at baseline[1].

The contribution of each component of the composite SSE between the arms will be evaluated. Descriptive statistics will be presented.

Sensitivity analyses will be conducted for the following scenarios[2]:

- SSE-FS will be evaluated using stratification factors derived from the information collected on the eCRF.

6.2.2 Secondary efficacy variables

The secondary efficacy variables are specified below:

- Overall survival (OS)
- Time to opiate use for cancer pain
- Time to pain progression
- Pain improvement rate
- Time to cytotoxic chemotherapy
- rPFS

6.2.2.1 Definition

Overall survival is defined as the time from the date of randomization to the date of death due to any cause. The OS time for subjects alive at the time of analysis (or database cut-off date) will be censored at their last known alive date. For subjects whose last date of follow-up confirms the

[1] As of Amendment 1.
[2] As of Amendment 3.

subject was alive at the data cut-off (i.e., based on the formal survival sweep), the subject will be censored at the data cut-off date. If a subject is lost to follow up and there was no contact after randomization, this subject will be censored at randomization date (OS = 1 day).

Time to opiate use for cancer pain is defined as the interval from the date of radium-223 dichloride treatment start to the date of start of opiate use. Subjects who have no opiate use at the time of analysis (or database cut-off date) will be censored at the last opiate use assessment date. Subjects with no on study assessment or no baseline assessment will be censored at radium-223 dichloride treatment start date (day 1). Subjects who have opiate use prior to radium-223 dichloride treatment start will be censored at radium-223 dichloride treatment start date. Time to opiate use for cancer pain will be analyzed based on the information collected on the analgesic concomitant medication; analgesics use 24 hour page and opiate use eCRF pages, whichever is the earliest.

Time to pain progression is defined as the interval from radium-223 dichloride treatment start to the first date a subject experiences pain progression based on worst pain subscale (WPS) and analgesic use. Time to pain progression will be evaluated in subjects with baseline WPS ≤ 8. Pain progression is defined as an increase of 2 or more points in the "worst pain in 24 hours" score from baseline observed at 2 consecutive evaluations $\geq$4 weeks apart

OR

An increase in pain management with respect to baseline, whichever occurs first.

An increase in pain management (IPM) is defined as.

- For subjects taking no analgesics or a non-opioid at baseline, the initiation of any opioid would be considered an IPM.
- For subjects taking a weak opioid at baseline, the initiation of any strong opioid would be considered an IPM.
- For subjects taking a strong opioid at baseline, the initiation of an additional strong opioid would be considered an IPM
- For subjects taking a weak opioid at baseline, the initiation of an additional weak opioid would be considered an IPM.[1]

Pain intensity (worst pain score [WPS]) assessments will occur on the day of visit. Subjects who have not experienced pain progression at the time of analysis (or database cut-off date) will be censored at the last post-baseline pain assessment date the subject was known to have not progressed. Subjects with no on-study assessment or no baseline assessment will be censored at radium-223 dichloride treatment start date[2].

[1] As of Amendment 3.
[2] As of Amendment 1.

The increase in pain management is based on the analgesic use information collected on all eCRF pages including analgesic concomitant medication, 24-hour analgesic use page and opiate use eCRF page.

Pain improvement is defined for subjects evaluable for pain improvement, i.e. subjects with baseline WPS≥ 2, as a 2-point decrease or more in WPS from baseline over 2 consecutive measurements conducted at least 4 weeks apart, without an increase in pain management.

Pain improvement rate is defined as the number of subjects with pain improvement as defined above, divided by the total number of subjects evaluable for pain improvement (i.e., subjects with baseline WPS≥ 2).

Time to cytotoxic chemotherapy is defined as the time from the date of randomization to the date of the first cytotoxic chemotherapy during follow up. Subjects who have not started cytotoxic chemotherapy during the study will be censored at last assessment for systemic anti-cancer therapy.

Radiological progression-free survival (rPFS) is defined as the time from the date of randomization to the date of first radiological progression or death (if death occurs before progression). For subjects without documented radiological progression or death at the time of analysis (or database cut-off date), i.e., data cut-off, the rPFS time will be censored at the date of the last evaluable tumor assessment. Every effort will be made to obtain radiologic scans for documentation of progression.

The radiological progression will be derived using algorithm in Table 5: Overall tumor response in appendix. Table 2 shows the end date and censoring rules for rPFS.

Subjects with baseline superscan will continue to have post-baseline bone imaging by either anatomically limited or full body CT/MRI, and the CT/MRI imaging will be used for rPFS analysis for these subjects. If no confirmed radiological progression is detected by bone imaging full-body CT/MRI at post-baseline and no death, then these subjects will be censored at the last full-body CT/MRI assessment date with clear radiological non-PD. Anatomically limited CT/MRI assessments identify results unknown.

Table 2: Radiological progression free survival censoring rules

Situation	End Date	Censored	Reason for Censoring
Subject had a radiological assessment of PD (no missed radiological assessment)	Date of first PD	No	N/A
Death during the study (no missed radiological assessment) or before first radiological PD assessment	Date of death	No	N/A
Subjects with baseline superscan, then either radiological PD(regardless by anatomically limited or full body CT/MRI) or death at post-baseline	Date of first PD or death	No	N/A
Subjects with baseline superscan, then radiological non-PD (by full body CT/MRI) and no death at post-baseline	Date of last full-body CT/MRI assessment with radiological non-PD(by full body CT/MRI) Or Date of randomization if no full body CT/MRI at post-baseline	Yes[1]	Radiological non-PD (by full body CT/MRI) and no death at post-baseline
Subjects without any evaluable post-baseline radiological tumor assessment.	The randomization date	Yes	No baseline or post-baseline tumor assessment.
Subjects who had a death or disease progression immediately after two or more consecutive missed radiological assessments.	Date of last radiological assessment before missed assessments.	Yes	Death or disease progression immediately after two or more consecutive missed radiological assessments.
New protocol prohibited systemic anticancer treatment started prior to radiological PD or death	Date of last radiological assessment prior to the change of therapy	Yes	New protocol prohibited systemic anticancer treatment started prior to radiological PD or death
Subjects who discontinue or withdraw early from the study without documented radiological disease progression.	Date of last evaluable radiological assessment	Yes	Discontinue or withdraw early from the study without documented radiological disease progression

Radiological Progression Free Survival (rPFS) = End Date – Start Date +1

Two consecutive missing radiological assessments are defined if the time interval between two consecutive radiological assessments is more than 26 weeks (2 × (12 +1) week) [2].

Table 2: Radiological progression free survival censoring rules

For patients with baseline superscan, radiological non-PD has to be identified by full-body CT/MRI. Anatomically limited CT/MRI assessments identify results unknown.

The prohibited concomitant therapies during the treatment phase of study include:

- Chemotherapy
- Radiopharmaceuticals, such as strontium-89, samarium-153, rhenium-186, or rhenium-188
- Hemibody external radiotherapy
- Other investigational drugs

6.2.2.2 Analysis of secondary efficacy endpoints

Time to event efficacy analysis will be performed to compare two treatment groups using a stratified log-rank test stratified by the same stratification factors as randomization. The hazard ratio (radium-223 dichloride / placebo) and its 80% and 95% CIs will be estimated from the Cox model, stratified by the same factors as for SSE-FS analysis. No alpha adjustment will be applied for secondary endpoints.

The product-limit estimates of the survival distribution functions (Kaplan-Meier) will be presented for each treatment group: N, total censored, total failed, time to event (median and its 80% CI, 25th and 75th percentile, range), probability of event-free survival at pre-specified months, such as 6, 12, 18, etc. months for OS. A Kaplan-Meier curve will be generated for each treatment group.

The pain improvement rates will be calculated by treatment group with their 80% CI as well as the differences of pain improvement rates between treatment groups and the corresponding 80% CIs. Pain improvement rate at week 12, EOT and any visit will be considered.

The comparison of pain improvement rate for all visits between the two treatment groups will be done using the “general association Cochran-Mantel-Haenszel statistic” stratified by the same stratification factors used in the SSE-FS analysis.

Due to inaccuracies in German translation of BPI-SF, the primary analyses of pain progression and pain improvement rates will be conducted with cleaned BPI-SF data only. Sensitivity analyses will be conducted for pain progression and pain improvement rates with all BPI-SF data, including uncleaned BPI-SF data.[3]

[1] As of Amendment 1.
[2] As of Amendment 1.
[3] As of Amendment 3.

6.2.3 Exploratory efficacy variables

The exploratory efficacy variables are specified below:

- Time to first on-study SSE
- Time to bone ALP progression
- Bone ALP response at week 12 and EOT
- Bone specific rPFS
- Time to visceral metastases onset

6.2.3.1 Definition

Time to first on-study SSE is defined as the time from the date of randomization to the date of the first on-study SSE. For subjects without SSE at the time of analysis, time to first on-study SSE will be censored at the date of last assessment of SSE.

The conventions for calculation of time to first on-study SSE are the same as for SSE-FS in section 6.2.1.1 except without considering death as an SSE event. The censoring rules for time to first on-study SSE are summarized in Table 3.

Table 3: Time to first on-study SSE censoring rule

Situation	End Date	Censored	Reason for Censoring
No post-baseline SSE assessment	Randomization date	Yes	No post-baseline SSE assessment
Subject had an SSE event	Date of first SSE	No	N/A
No SSE and no death at data cutoff	Last SSE assessment	Yes	No SSE
No SSE prior to death	Last SSE assessment	Yes	No SSE prior to death

SSE events immediately after missing SSE assessments are still counted as events in the analysis of Time to first on-study SSE.

Time to bone ALP progression is defined as the time from the date of randomization to the date of first bone ALP progression. Bone ALP progression is defined as a ≥25% increase from the baseline value, at least 12 weeks from baseline in subjects with no bone ALP decline from baseline; or a ≥25% increase above the nadir value, which is confirmed by a second value obtained 3 or more weeks later in subjects with an initial bone ALP decline from baseline. Confirmation visit does not have to be adjacent to first visit with increase from nadir. Subjects without bone ALP progression at the time of analysis (or database cut-off date) will be censored at last bone ALP assessment. Subjects without baseline or post-baseline bone ALP values will be censored at date of randomization.

For subjects who had a bone ALP progression immediately after two or more consecutive missed ALP assessments, time to bone ALP progression will be censored at the date of last bone ALP assessment before missed assessments. Two consecutive missing assessments is defined if the time interval between two consecutive ALP assessments is more than 10 weeks (2 × (4 weeks +1

week))during the treatment period, or 26 weeks (2 × (12 weeks +1 week)) during the active follow-up period[1].

Bone ALP response at week 12 and EOT is defined as a ≥30% reduction of the blood level at Week 12 and EOT, compared to the baseline value. Confirmed bone ALP response is defined as a ≥30% reduction of the blood level, compared to the baseline value, confirmed by a second bone ALP value 4 or more weeks later.

Bone specific rPFS is defined as the time from the date of randomization to the date of confirmed radiological progression detected by bone imaging or death (if death occurs before progression). Subjects without confirmed radiological progression or death at the time of analysis (or database cut-off date) will be censored at their last evaluable radiological **bone** imaging assessment. If a new bone lesion or unequivocal increase in size of bone lesions is identified on bone scan, the lesion must also be confirmed by computed tomography/magnetic resonance imaging (CT/MRI). If a new bone lesion or unequivocal increase in size of bone lesions is only visible on a CT/MRI and not visible on a technetium 99m bone scan, progression will be declared without further confirmation. The conventions for calculation of bone specific rPFS are the same as for rPFS in section 6.2.2.1 except that **bone** progression or death (if death occurs before progression) is considered as a bone specific rPFS event and **bone** tumor assessment is used to determine the event/censoring.
Subjects with baseline superscan will continue to have post-baseline bone imaging by either anatomically limited or full body CT/MRI, and the CT/MRI imaging will be used for bone specific rPFS analysis for these subjects. If no confirmed radiological progression is detected by bone imaging full-body CT/MRI at post-baseline, then these baseline superscan subjects will be censored at the last full-body CT/MRI assessment date with clear radiological non-PD. Baseline superscan subjects will be censored at date of randomization if no full body CT/MRI at post-baseline. Anatomically limited CT/MRI assessments identify results unknown.

Time to visceral metastases onset is defined as the time from the date of randomization to the date of the first scan showing visceral metastatic disease.

The censoring rules for time to visceral metastases onset are summarized in Table 4.

Table 4: Time to first visceral metastases onset censoring rule

Situation	End Date	Censored	Reason for Censoring
Subject had a visceral metastases	Date of the first scan showing visceral metastatic disease	No	N/A
Subjects without any evaluable post-baseline non-bone tumor assessment.	The randomization date	Yes	No baseline or post-baseline tumor assessment.
Subjects who didn't have visceral metastases at the time of analysis.	Date of last evaluable non-bone tumor assessment	Yes	Death or disease progression immediately after two or more consecutive missed

[1] As of Amendment 3.

			radiological assessments.

6.2.3.2 Analysis of exploratory efficacy endpoints

Similar to the primary efficacy endpoint, time to event efficacy analysis will be performed to compare two treatment groups using a stratified log-rank test stratified by the same stratification factors as randomization. The hazard ratio (radium-223 dichloride / placebo) and its 80% and 95% CIs will be estimated from the Cox model, stratified by the same factors as for SSE-FS analysis.

The product-limit estimates of the survival distribution functions (Kaplan-Meier) will be presented for each treatment group: N, total censored, total failed, time to event (median and its 80% CI, 25th and 75th percentile, range), rate of subjects who did not yet develop an event at pre-specified months. A Kaplan-Meier curve will be generated for each treatment group.

Sensitivity analysis of Time to bone ALP progression will be performed by dropping the 12 week restriction in the definition.

The bone ALP response at Week 12 and EOT will be calculated by treatment group with their 80% CI as well as the differences of bone ALP response between treatment groups and the corresponding 80% CIs.

The comparison of bone ALP response at Week 12 and EOT between two treatment groups will be done using the "general association Cochran-Mantel-Haenszel statistic" stratified by the same stratification factors used in the SSE-FS analysis.

The change from baseline WPS to the Week 12 visit and EOT visit will be calculated. The missing values at Week 12 and /or EOT will remain missing (the previous value will not be carried forward).

6.2.4 Subgroup analyses

Descriptive statistics and hazard ratio estimates with 80% CIs will be provided for SSE-FS and rPFS at least within each category of the following variables, provided there are a sufficient number of events in total within the subgroup across the treatment groups.

The list used for subgroup analyses for SSE-FS were updated and provided as below:

- Age (≥55, <55) [1]
- Race
- Geographical region (Europe/North America [including Israel] versus Asia)

[1] As of Amendment 1.

- Baseline ECOG performance status
- Previous lines of hormone therapy in metastatic setting
- Visceral disease[1]
- Prior SREs
- Baseline total body weight
- Baseline ideal body weight
- BMI.

The subgroup analyses for rPFS will be performed as below[2]:

- Age (≥55, <55)
- Geographical region (Europe/North America [including Israel] versus Asia)
- Baseline ECOG performance status
- Previous lines of hormone therapy in metastatic setting
- Visceral disease
- BMI

6.2.5 Brief Pain Inventory-Short Form (BPI-SF)

The BPI-SF is an 11-item, self-administered, clinically valid, reliable, and responsive measure developed to assess pain related to cancer. All BPI items are scored using rating scales. Four items measure pain intensity (pain now, average pain, worst pain, and least pain) using 0 (no pain) to 10 (pain as bad as you can imagine) numeric rating scales, and 7 items measure the level of interference with function caused by pain (general activity, mood, walking ability, normal work, relations with other people, sleep, and enjoyment of life) using 0 (no interference) to 10 (complete interference) rating scales.

The items are aggregated into 2 dimensions: (1) Pain severity index, and (2) Function interference index, using the mean of the 7 pain interference items. All 4 severity items must be completed for aggregating the pain severity index. The function interference index is scored as the mean of the item scores multiplied by 7, given that more than 50% or 4 of 7 of the items have been completed.

Descriptive statistics for observed data will be presented at each assessment time point and for change from baseline by treatment group.

[1] As of Amendment 1.
[2] As of Amendment 3.

6.2.6 Other variables

6.2.6.1 Resource utilization

Resource utilization data will be presented in the listing.[1]

6.3 Safety Analysis

The summaries of the safety data will be completed for the Safety analysis population (SAF). No formal statistical test will be performed for the safety variables.

6.3.1 Extent of exposure

Study medication and single standard of care hormonal treatment will be summarized for the SAF by treatment group, using descriptive statistics such as frequency and proportion (for categorical variables), mean, median, and standard deviation (for continuous variables).

Summaries will include, but may not be limited to:

- Duration of treatment Radium-223 dichloride: overall duration of study treatment, number of injections of study treatment, time between two consecutive injections, such as time between injections X and X+1, X=1, 2, …, 5, average time between injections. Overall duration of single standard of care hormonal treatment in study medication treatment period and follow up period, respectively.
- Dosage of Radium-223 dichloride: total activity injected, total activity injected per body weight. Dosage of hormonal treatment will be summarized for safety population.
- Dose modifications of Radium-223 dichloride: number of subjects with all injections performed successfully and without delay, number of injections not performed successfully, number of injections delay, reason for injections delay. Dose modification of hormonal treatment.

6.3.2 Adverse events

All adverse events (AEs) will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA) Version 18 or most recent version. The intensity of an AE will be documented using National Cancer Institute (NCI) Common Terminology Criteria Adverse Event (CTCAE), version 4.03.

A treatment-emergent AE is defined as any event arising or worsening after the start of study drug administration until 30 days after the last study medication intake.

[1] As of Amendment 3

Summary statistics (frequency and percentage of subjects, not of events) will be presented by treatment group using MedDRA by worst grade for the following:

- Incidence rate of pre-treatment AEs.
- Incidence rate of treatment-emergent AEs.
- Incidence rate of treatment-emergent drug-related AEs.
- Incidence rate of treatment-emergent AEs with grades 3 or 4.
- Incidence rate of treatment-emergent AEs with Grade 5.
- Incidence rate of treatment-emergent AEs leading to permanent withdrawal of medication.
- Incidence rate of treatment-emergent AEs leading to dose interruptions.
- Incidence rate of treatment-emergent AEs leading to dose reductions (applies to background treatment and concomitant medication of interest).
- Incidence rate of serious treatment-emergent adverse events.
- Incidence rate of serious treatment-emergent drug-related adverse events.
- Listing of treatment-emergent AEs leading to withdrawal: subject ID, investigator AE term, start and stop date of study drug administration, start and stop date of AE, drug related (yes/no), serious (yes/no), worst grade, outcome.
- Listing of treatment-emergent serious AEs: subject ID, investigator AE term, worst grade, start and stop dates of study treatment, start and stop date of AE , drug related (yes/no), outcome, action taken.
- In addition, descriptive summaries will be provided for long term safety endpoints including new primary malignancies and hematopoietic reserve for tolerability of subsequent chemotherapy.
- Bone fractures and bone associated events will be reported, including during long-term follow-up, regardless of the investigator's causality assessment.[1]

6.3.3 Deaths

Deaths reported during the study period will be tabulated by treatment group.

- Summary table of deaths (all deaths, all deaths during treatment and up to 30 days after last dose of study drug, all deaths later than 30 days after last dose of study medication)

[1] As of Amendment 2.

- Listing of subjects who died during treatment and up to 30 days after last dose: subject ID, start and stop date of study medication, date of death, and cause of death.

6.3.4 Clinical laboratory evaluations

Worst grades for hematological and biochemical toxicities will be calculated according to CTCAE, version 4.03 based on laboratory measurements. Summary statistics (frequency and percentage of subjects, not of events) will be presented by treatment group and NCI CTC worst grade for the following:

- Hematological and Biochemical Toxicities: treatment-emergent events by worst CTCAE grade.
- Change in worst grade from baseline for hematological and biochemical toxicity (worst grade under treatment- latest pre-treatment value).
- Shift table from baseline to worst grade post-baseline for hematological and biochemical toxicity (worst grade under treatment- latest pre-treatment value).
- By-subject listing of subjects with abnormal laboratory values.

If more than one assessment occurred at any post-baseline visit (repeated measures at same visit), 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. CTC grading of some laboratory parameters require clinical information in addition to the laboratory values. However, laboratory CTC grades presented in the CSR will be based only on the numerical lab results.

6.3.5 Other safety measures

The last pre-treatment safety measurement, i.e., SBP, DBP, weight, body temperature, heart rate, respiration rate and electrocardiogram (ECG) will be used as “baseline value.”

When more than one value is collected at the same visit, the value retained at that particular visit for summary statistics will be the average of the different measures reported for that visit.

For each treatment group, vital signs will be tabulated and summarized by visit for observed values and changes from baseline using descriptive statistics, as appropriate.

Summary statistics of ECG data will be reported at each visit for raw data and for changes from baseline. In addition, incidence rates of treatment-emergent ECG findings will be reported.

The number and percentage of subjects in each ECOG category will be presented by visit. Changes from baseline in ECOG will be summarized in shift tables by treatment group and visit.

6.4 Pharmacokinetics

No pharmacokinetic (PK) measurements will be performed in this study.

6.5 Biomarker analyses

Biomarker analyses planned within this study may include predictive, prognostic, and pharmacodynamic biomarkers analyzed from serum and urine. Summary statistics for biomarkers and their changes from baseline may be presented by visits and treatment depending on the availability of data.

Biomarker analyses will be described in a separate analysis plan.

6.6 Sample size estimation

The sample size is based on the primary efficacy endpoint, SSE-FS.

Assuming 1-sided alpha of 0.1 (2-sided alpha of 0.2), power of 90%, and a randomization ratio of 1:1 between the experimental and control arms, 119 events are required to detect a 60% increase in SSE-FS. The expected study duration is 24.0 months assuming subjects enroll at a rate of 20 subjects per month, an enrollment ramp-up time of 6 months, a dropout rate of 15% for the primary endpoint, exponentially distributed event time, 8.3 month median time for the control group and a 14.8 month long enrollment for a total of 227 patients in the two treatment groups combined.

6.7 Additional analyses planned to be reported outside the main report

Biomarker analyses if any will be reported in a separate report.

6.8 Change of analysis[1]

6.8.1 Efficacy endpoints

The planned sensitivity analyses for all efficacy endpoints in previous sections will not be performed due to the study early termination, except for SSE-FS will be evaluated using stratification factors derived from the information collected on the eCRF. All confidence intervals for time to event results specificied in previous sections are calculated for 80% CI only. A hazard ratio with 2-sided 80% and 95% CI from a stratified Cox proportional hazards model will be performed as planned.

The pain improvement rate will be summarized, however, the planned comparison of pain improvement rate using the “general association Cochran-Mantel-Haenszel statistic” will not be performed.

[1] As of Amendment 2.

The list used for subgroup analyses for SSE-FS were updated and provided as below:

- Age (≥55, <55)
- Race
- Geographical region (Europe/North America [including Israel] versus Asia)
- Baseline ECOG performance status
- Previous lines of hormone therapy in metastatic setting
- Visceral disease
- Prior SREs,
- Baseline total body weight
- Baseline ideal body weight
- BMI

The subgroup analyses for rPFS will be performed as below:

- Age (≥55, <55)
- Geographical region (Europe/North America [including Israel] versus Asia)
- Baseline ECOG performance status
- Previous lines of hormone therapy in metastatic setting
- Visceral disease
- BMI

6.8.2 Other endpoints

The following planned analysis will not be performed due to the study early termination.

- Resource utilization (will be included in listing only)
- Biomarker analyses
- Area under curve (AUC) for pain severity index and function interference index

7. Document history and changes in the planned statistical analysis

- Signed SAP version 1.0 dated 29 APR 2016, according to protocol version 1.0 up to amendment 5, integrated protocol version 5.0, and dated 11 MAR 2016.

- Signed SAP version 2.0 dated 25 OCT 2017, according to protocol version 1.0 up to amendment 7, integrated protocol version 6.0, and dated 23 MAY 2017.

- Signed SAP version 3.0 dated 04 JAN 2019, according to protocol version 1.0 up to amendment 8, integrated protocol version 8.0, and dated 03 APR 2018.

- Signed SAP version 4.0 dated 05 AUG 2019, according to protocol version 1.0 up to amendment 8, integrated protocol version 8.0, and dated 03 APR 2018.

8. References

1. Rubens RD, Coleman RE. In: Abeloff MD, Armitage JO, Lichter AS, Niederhuber JE. Clinical oncology. New York: Churchill Livingstone, 1995:643-65.
2. Cessna JT, Zimmerman BE. Standardization of radium-223 by liquid scintillation counting. Appl Radiat Isot. 2010;68(7-8):1523-8.
3. Zimmerman BE, Bergeron DE, Cessna JT, Fitzgerald R, Pibida L. Revision of the NIST standard for 223Ra: new measurements and review of 2008 data. Journal of Research of the National Institute of Standards and Technology. 2015;120:37-57.

9. Appendices

9.1 Overall tumor response

Table 5: Overall tumor response

Overall non-bone response according to modified RECIST 1.1	Response of non-target bone lesions (including new bone lesions)	Overall response (consider both non-bone and bone response)	The scan date used for censoring if overall response is non-PD (1st or last date)
Complete response	Complete response	CR	
Partial response	Complete response	PR	
Stable disease	Complete response	SD	
Non CR/Non PD	Complete response	Non CR/Non PD	
Progressive disease	Complete response	PD	Last assessment date prior to PD
Not evaluable/Missing	Complete response	NE	
Complete response	Non CR/Non PD	PR	
Partial response	Non CR/Non PD	PR	
Stable disease	Non CR/Non PD	SD	
Non CR/Non PD	Non CR/Non PD	Non CR/Non PD	
Progressive disease	Non CR/Non PD	PD	Last assessment date prior to PD
Not evaluable/Missing	Non CR/Non PD	NE	
Complete response	Unequivocal progression	PD	Bone response date
Partial response	Unequivocal progression	PD	Bone response date
Stable disease	Unequivocal progression	PD	Bone response date
Non CR/Non PD	Unequivocal progression	PD	Bone response date
Progressive disease	Unequivocal progression	PD	Bone response date
Not evaluable/Missing	Unequivocal progression	PD	Bone response date
Complete response	Not (all) evaluated/Missing	NE	

Table 5: Overall tumor response

Overall non-bone response according to modified RECIST 1.1	**Response of non-target bone lesions (including new bone lesions)**	**Overall response (consider both non-bone and bone response)**	**The scan date used for censoring if overall response is non-PD (1st or last date)**
Partial response	Not (all) evaluated/Missing	NE	
Stable disease	Not (all) evaluated/Missing	NE	
Non CR/Non PD	Not (all) evaluated/Missing	NE	
Progressive disease	Not (all) evaluated/Missing	PD	Last assessment date prior to PD
Not evaluable/Missing	Not (all) evaluated/Missing	NE/Missing	
Not applicable	Complete response	CR	Bone response date
Not applicable	Non CR/Non PD	Non CR/Non PD	Bone response date
Not applicable	Unequivocal progression	Unequivocal progression	Bone response date
Not applicable	Not (all) evaluated	Not (all) evaluated	Bone response date

Definitions:
1. Not applicable = No non bone lesions at baseline and no new non bone lesions (no PD)
2. Not evaluated = not all evaluated = Missing (new lesion must be absent)
3. As per RECIST 1.1 paper: ""When no imaging/measurement is done at all at a particular time point, the patient is not evaluable (NE) at that time point. If only a subset of lesion measurements are made at an assessment, usually the case is also considered NE at that time point"" (unless other arguments for PD)"