Document Type:	Statistical Analysis Plan and supplement
Official Title:	A phase II randomized, double-blind, placebo-controlled trial of radium-223 dichloride in combination with exemestane and everolimus versus placebo in combination with exemestane and everolimus when administered to metastatic HER2 negative hormone receptor positive breast cancer subjects with bone metastases
NCT Number:	NCT02258451
Document Date:	14 Feb 2020 24 Jun 2020

Title Page

A phase II randomized, double-blind, placebo-controlled trial of radium-223 dichloride in combination with exemestane and everolimus versus placebo in combination with exemestane and everolimus when administered to metastatic HER2 negative hormone receptor positive breast cancer subjects with bone metastases

Short title: Study of radium-223 dichloride in combination with exemestane and everolimus versus placebo in combination with exemestane and everolimus 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 in combination with exemestane and everolimus versus placebo in combination with exemestane and everolimus when administered to metastatic HER2 negative hormone receptor positive breast cancer subjects with bone metastases		
Clinical study phase:	II	**Date:**	14 FEB 2020
Study No.:	BAY 88-8223 / 17096	**Version:**	V4.0
Author:	PPD Bayer HealthCare Pharmaceuticals Whippany, NJ 07981		

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
EOXT	End of Xofigo 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
TEAE	Treatmemt-emergent adverse event
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

Study Background

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.

Recently, a phase III BOLERO-2 study showed that the addition of everolimus to exemestane significantly improved progression-free survival (PFS) compared with single-agent exemestane in estrogen receptor positive (ER+), HER2-negative patients whose metastatic disease was refractory to prior treatment with letrozole or anastrozole.[[3]] This led to the recent marketing approval of everolimus by the FDA and EMA for the treatment of hormone receptor-positive, HER2/neu negative advanced breast cancer, in combination with exemestane, in postmenopausal women without symptomatic visceral disease after recurrence or progression following a NSAI. At the time of final analysis with a median follow-up of 18 months, the median PFS (as per Investigator review) was 7.8 months for everolimus and exemestane versus 3.2 months for placebo and exemestane (hazard ratio [HR] = 0.45, 95% confidence interval [CI]: 0.38 to 0.54; log-rank $p<0.0001$).[[4]]

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

This Study

This study is a phase II, randomized, double-blind, placebo-controlled, parallel group trial of radium-223 dichloride versus placebo administered with exemestane and everolimus and supportive care in subjects with HER2 negative hormone receptor positive breast cancer with bone metastases. The primary endpoint is skeletal event-free survival. At study start the study had a single-stage design with no interim analysis. Protocol Amendment 8 added an unblinded administrative review, which was performed on 05 DEC 2017. Subsequent to the review, and prior to the primary analysis, Protocol Amendment 10 curtailed post-treatment follow-up and permitted patients to be transferred to a long-term follow-up study following their end-of-treatment visit.

This Statistical Analysis Plan

This statistical analysis plan (SAP) describes the analyses to be performed for the final primary analysis of study 17096, to be performed at primary endpoint maturation and included in the Clinical Study Report. In addition, it describes the analyses to be performed for the interim administrative review and for the exploratory updated analyses to be performed after study closure. This document is version 4.0 (amendment 3) of the SAP. It is based on the SAP version 1.0, dated 25 OCT 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 7.0

(amendment 10), dated 04 DEC 2019.[1] The original protocol, version 1.0, is dated 06 MAY 2014. Detailed descriptions of the Tables, Listings, and Figures (TLFs) to be created will be described in a separate TLF specification document.

This SAP excludes any biomarker analyses, which will be described in a separate document and reported in a separate report.

2. Study Objectives

The objective of this study is to assess efficacy and safety of radium 223 dichloride in combination with exemestane and everolimus in subjects with human epidermal growth factor receptor 2 (HER2) negative, hormone receptor positive breast cancer with bone metastases.

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

The exploratory endpoints are:

- time to first on-study SSE;
- time to bone-ALP progression;

[1] As of Amendment 3.

- bone-ALP response at Week 12 and 4 weeks (± 7 days) after last radium-223 dichloride/placebo dose;
- bone-specific rPFS;
- resource utilization;
- biomarker assessments,
- time to visceral metastases onset.

3. Study Design

This study is a phase II, randomized, double-blind, placebo-controlled, parallel group trial of radium-223 dichloride versus placebo administered with exemestane and everolimus and supportive care 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.
- Visceral disease: Yes versus No

This study will be conducted at approximately 160 investigative study sites and approximately 311 subjects will be enrolled.

Prior to the primary analysis, the study is comprised of 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[[5]][1]) (Arm A) or placebo (Arm B) each separated by an interval of 4 weeks. All subjects will start treatment with exemestane and

[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.

everolimus, after randomization, either before or simultaneously to the first injection of radium-223 dichloride/placebo. Treatment with exemestane and everolimus will continue after completion of radium-223 dichloride until disease progression, initiation of new anti-cancer therapy, unacceptable toxicity occurs , discontinuation of study treatment for other reasons, or study is terminated, whichever occurs first. All subjects receive supportive care, as per local standard of practice.

During the treatment period, subjects will be assessed for safety at each treatment visit, every 8 weeks, and will be evaluated for radiological progression according to local standard practice. Symptomatic skeletal events should be recorded until end of treatment visit, independent of whether the subject starts a new anti-cancer therapy (i.e., chemotherapy, other).Subjects will receive an end-of treatment clinic safety visit 30 days +7 following their last treatment.

Prior to Protocol Amendment 9, subjects who discontinued or completed the treatment phase will enter an active follow-up phase. Following Protocol Amendment 9, the active follow-up phase became optional. For subjects who have not experienced an SSE. active follow-up visits will occur every 4 weeks ± 7 days, with clinic visits where feasible, otherwise by telephone. For subjects who have experienced an SSE, active follow-up will occur by telephone every 8 weeks ± 7 days. Per Protocol Amendment 10, the active follow-up phase will be discontinued entirely following the cutoff for the primary analysis.

This study will end, and the updated analysis will be performed, after the last 4 weeks follow up visit by the time last subject on treatment discontinues oral exemestane and/ or everolimus treatment and completes end of treatment visit or discontinues from this study for another reason (e.g., death, consent withdrawn and active objection for further data collection, lost to follow-up).

Changes to Study Design

Protocol Amendment 8, dated 18 May 2017, provided for an unblinded administrative interim data review of efficacy and safety, with efficacy focusing on rPFS, to be performed when a minimum of 80 rPFS events is reached. The administrative data review took place on 5 Dec 2017.

Protocol Amendment 9, dated 3 Apr 2018, permitted subjects who completed the later of end of treatment visit or 30 days from last study treatment to transfer into a separate long-term follow-up study, whether or not an SSE had been observed. In addition, enrollment was discontinued in April 2018.

Protocol Amendment 10, dated 4 Dec 2019, removed the active follow-up period from this study following the primary analysis. After the primary analysis, subjects who complete the later of end of treatment visit or 30 days from last study treatment will be required to transfer into a separate long-term follow-up study.

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

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 continuous 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, with more specific standards controlling over more global ones. Where needed for study-specific reasons, study-specific departures from standards may be specified

4.2 Handling of Withdrawal

A "withdrawal" from treatment is defined as a subject who has been randomized and discontinues study treatment 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.

A "screen failure" is defined as a subject who enrolled in the study but was not randomized.

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.

Subjects who were lost to follow-up will be handled in time to event analyses per the applicable censoring rules.

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.

No imputation will be performed for missing lesion assessment and tumor response, unless otherwise specified.

The partial missing date used for derivation of time to event will be imputed and the imputation rule will be specified in the analysis dataset specifications.

Specific statistical methods, such as handling of censoring for time-to-event methods (e.g. Kaplan-Meier and the log-rank test), address loss to follow-up and other missing data situations through statistical assumptions established in the literature.

4.4 Data Rules

Safety baseline value is defined as the last non-missing value on or prior to first administration date of any study medication (Radium-223 dichloride/placebo or exemestane or everolimus).

Investigator determine the baseline for tumor assessments based on protocol rules. For other efficacy endpoints except pain endpoints, the baseline is defined as the last available value prior to or on the date of randomization. For pain efficacy endpoints baseline is defined as last available value prior to first administration of radium-223 dichloride/ Placebo since the data are not collected at screening visit.

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. Interim Analyses and Data Monitoring

4.4.1 Interim Analyses

Protocol Amendment 8, dated 18 May2017, added an administrative interim data review, to be performed when a minimum of 80 rPFS events is reached. The interim data review is to be primarily focused on the rPFS, with the interim data review results to inform future radium-223 clinical development plans in this indication. The review was to be conducted by an independent unblinded review committee..

No formal statistical testing was specified for either SSE-FS or rPFS at the time of this interim review. Both summary statistics and hazard ratio 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. The hazard ratio (radium-223 dichloride / placebo) for rPFS and its 80% and 95% confidence intervals (CIs) will be calculated using a univariate Cox model, stratified by stratification factors: geographical regions (Europe/North America [including Israel] vs. Asia), previous lines of hormone therapy (1 vs. ≥2), and visceral disease (yes vs. no). Sensitivity analyses will be conducted for rPFS with Cox model stratified by prior SREs (1 vs. 2), visceral disease (yes vs. no), and non-stratified Cox model, respectively.

No adjustment to the original single-stage operating characteristics is planned as a result of this interim analyis. As there was no plan to stop the trial due to superior efficacy, no alpha adjustment is applied for this interim look at the primary endpoint of SSE-FS. In addition, as at the time this interim analysis was planned, there was no plan to stop the study for futility, no beta

spending or power adjustment is planned. As an interim analysis may nonetheless affect study operating characteristics, some caution in interpretation may be warranted. The study will remain blinded after the interim data review. [1]

The planned interim efficacy analysis focused on rPFS and included summary statistics, Kaplan-Meier estimates, and hazard ratios with confidence intervals for rPFS and SSE-FS. The planned interim safety analysis included treatment-related AEs, study drug-related TEAEs, TE, AEs of interest, serious AEs, and lab toxicities.Details are described in the separate administrative review analysis specifications document.[2]

4.4.2 Data Monitoring

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

4.5 Primary and updated Analyses

After implementation of **CSP Amendment 9**, subjects who completed the EOT visit could be transferred to a separate extended safety follow-up study for their remaining follow-up, whether or not they had experienced an SSE or radiological progression. In face of the enrollment discontinuation on Apr/2018 and potentially curtailed active follow-up introduced by **CSP Amendment 9,** the sponsor decided to conduct the primary analysis, independent of the number of SSE-FS achieved, with a cutoff date of 22JAN2020. Therefore, there will be 2 analyses in the study: the primary analysis will encompass all data from study start until cutoff date of 22JAN2020, and an exploratory updated analysis (after implementation of **CSP Amendment 10**) will be performed after the last subject discontinues treatment discontinues oral exemestane and/or everolimus treatment and completes end of treatment visit or discontinues from this study for another reason (e.g., death, consent withdrawn and active objection for further data collection, lost to follow-up) .

4.5.1 Primary Analysis

Following the enrollment discontinuation on Apr/2018 and observation of curtailed active follow-up introduced by **CSP Amendment 9,** the sponsor decided to conduct the primary analysis, independently of the number of SSE-FS achieved, with a cutoff date of 22JAN2020. This analysis will include all of the items described in Section 6 below and include all of the planned tables, figures, and listings for this study. It will include all data from the treatment period for each subject and all follow-up data up until the date that the last subject who completes his End of Treatment visit are transferred to a separate extended safety follow-up study.

---

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

4.5.2 Updated Analysis

A follow-up analysis will be conducted when all treated subjects discontinue oral treatment and are transferred to or decline the separate extended safety follow-up study. This analysis will include the following items:

- Disposition;
- Adverse events: related adverse events, related serious adverse events (including deaths), new primary malignancies, new symptomatic skeletal event-related adverse events, deaths and adverse events of interest;
- Efficacy: symptomatic skeletal event free-survival, overall survival.

4.6 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

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 safety analysis (SAF) population is comprised of all randomized subjects who received at least one dose of any study medication (Radium-223 dichloride or placebo or exemestane or everolimus) 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, 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), Nottingham combined histologic grading 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)), 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

[1] As of Amendment 1.

versus 2 or more), and prior SREs (1 versus 2), visceral disease (yes versus no), cancer pain assessment

6.1.3 Medical history

The dictionary for coding is MedDRA version 19 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.
- 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.

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.

To demonstrate SSEs, one or more associated symptoms need to be recorded on the applicable AE CRF and linked to an associated procedure (e.g. EBRT, imaging, surgery, etc.). For event date, the procedure date instead of the dates of symptom(s) will be used to calculate the SSE-FS.

The censoring rules for SSE-FS are summarized in Table 6–1.

Table 6–1: 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:

Per protocol, the primary analysis of SSE-FS was to be performed when approximately 160 subjects have SSE-FS.

Following the enrollment discontinuation on APR 2018 and observation of curtailed active follow-up introduced by CSP Amendment 9, the sponsor decided to conduct the primary analysis, independently of the number of SSE-FS achieved, with a cutoff date of 22JAN2020.

Since differences are expected between the values of the stratification variables entered by the investigator at the time of randomization (IxRS) 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 IxRS. Stratification factors collected on the eCRF will be used for sensitivity analyses. In rare circumstances, if eCRF data are not available, then IxRS data will be used.

The null hypothesis that both treatment with radium-223 dichloride dose not result in superior SSE-FS to treatment with placebo in subject population was tested against the alternative hypothesis that the treatment with radium-223 dichloride results in superior SSE-FS time to treatment the placebo.

H_0: $\text{SSE-FS}_{\text{Radium-223+ Exemestane/Everolimus}} \leq \text{SSE-FS}_{\text{ Placebo+ Exemestane/Everolimus,}}$ versus

H_A: $\text{SSE-FS}_{\text{Radium-223+ Exemestane/Everolimus}} > \text{SSE-FS}_{\text{ Placebo+ Exemestane/Everolimus}}$

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 visceral disease, based on the stratification data in IxRS.

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 95% CI, 25^{th} and 75^{th} percentile, range), probability of event-free survival 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 95% CIs will also be calculated by treatment group.

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[1] if the primary analysis of SSE-FS result is significant:

- SSE-FS will be re-evaluated considering initiation of anti-cancer therapies (excluding standard of care anti-cancer hormonal treatments). SSE-FS will be censored at the last SSE assessment before the start of new anti-cancer therapies (excluding standard of care anti-cancer hormonal treatments).

- SSE-FS will be evaluated considering all deaths as events, regardless duration of 9 weeks between last SSE assessment and death.

- SSE-FS will be evaluated with an unstratified log-rank testSSE-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 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.

---

[1] As of Amendment 3.

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 ≥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 . Subjects with no on-study assessment or no baseline assessment will be censored at radium-223 dichloride treatment start date.

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.

---

[1] As of Amendment 3.

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), the rPFS time will be censored at the date of the last evaluable tumor assessment. Every effort will be made to obtain radiological scans for documentation of progression.

The radiological progression will be derived using algorithm in Table 9–1: Overall tumor response in appendix. Table 6–2 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 6–2: Radiological progression-free survival (rPFS) 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 new systemic anticancer treatment* started prior to radiological PD or death	Date of last radiological assessment prior to the change of therapy	Yes	New protocol new systemic anticancer treatment* started prior to radiological PD or death

[1] As of SAP Amendment 1.

Table 6–2: Radiological progression-free survival (rPFS) censoring rules

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 18 weeks (2 × (8 +1 week)). [1]

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

* **E**xcluding standard of care anti-cancer hormonal treatments

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
- All medications that are prohibited as per the local label instructions for exemestane and everolimus and the supportive treatment.

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 95% CIs will be estimated from the Cox model, stratified by the same factors as for SSE-FS analysis. No alpha adjustment for multiplicity 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 95% 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.

[1] As of SAP Amendment 1.

The pain improvement rates will be calculated by treatment group with their exact binomial Clopper-Pearson 95% CI as well as the differences of pain improvement rates between treatment groups and the corresponding exact binomial Clopper-Pearson 95% CIs. Pain improvement rate at week 12, EOXT and any visit will be considered.

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, excluding the mis-translated data from German sites. Sensitivity analyses will be conducted for pain progression and pain improvement rates with all BPI-SF data, including uncleaned BPI-SF data.[1]

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

- The following sensitivity analyses will be conducted for rPFS if the primary analysis of rPFS result is significant :rPFS will be evaluated with an unstratified log-rank test.

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

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 EOXT
- 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 6–3.

[1] As of Amendment 3.

Table 6–3: 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 without any 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 17 weeks (2 × 8 weeks +1 week) during the active follow-up period[1].

Bone ALP response at week 12 and EOXT is defined as a ≥30% reduction of the blood level at Week 12 and EOXT, 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

[1] As of Amendment 3.

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 (in subjects with no visceral disease at baseline) 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 6–4.

Table 6–4: Time to first visceral metastases onset censoring rule

Situation	End Date	Censored	Reason for Censoring
Subject had a visceral metastases at post-baseline	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 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	Subjects who didn't have visceral metastases at the time of analysis.

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 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 95% 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 EXOT will be calculated by treatment group with their exact binomial Clopper-Pearson 95% CI as well as the differences of bone ALP response between treatment groups and the corresponding exact binomial Clopper-Pearson 95% CIs.

The comparison of bone ALP response at Week 12 and EOXT 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 in WPS at Week 12 visit and EOXT visit will be calculated. The missing values at Week 12 and /or EOXT will remain missing (the previous value will not be carried forward).

6.2.4 Subgroup analyses

Descriptive statistics and hazard ratio estimates with 95% 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)
- 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
- Opiate use

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

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

- Previous lines of hormone therapy in metastatic setting
- Visceral disease
- BMI
- Baseline total body weight

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.

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.

[1] As of Amendment 3

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, average time between injections.
- Overall duration of exemestane and/or everolimus treatment in study medication treatment period, respectively and together.
- Dosage of Radium-223 dichloride: total activity injected, total activity injected per body weight.
- Actual dosage per day of exemestane and everolimus treatment will be summarized for safety population.
- Dose modifications of Radium-223 dichloride: number of injections delayed, number of injection delays due to AE, and reasons other than AE.
- Dose modification of exemestane and everolimus treatment: number of delays, number of delays due to AE, and reasons other than AE.

6.3.2 Adverse events

All adverse events (AEs) will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA) Version 19 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 started or worsen between first dose date of any study medication (Radium-223 dichloride/placebo or exemestane or everolimus) and last dose date of any study medication + 30 days, inclusive.

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

- Incidence rate of pre-treatment AEs.
- Incidence rate of treatment-emergent AEs.
- Incidence rate of treatment-emergent drug-related AEs, with respect to Radium-223 dichloride/placebo, Exemestane and Everolimus, respectively.
- Incidence rate of treatment-emergent AEs with grades 3 or 4.

- Incidence rate of treatment-emergent AEs leading to death.
- Incidence rate of treatment-emergent AEs leading to permanent withdrawal of medication of Radium-223 dichloride/placebo, Exemestane and Everolimus, respectively.
- Incidence rate of treatment-emergent AEs leading to dose interruptions of Radium-223 dichloride/placebo, Exemestane and Everolimus, respectively.
- Incidence rate of treatment-emergent AEs leading to dose reductions (applies to background treatment and concomitant medication of interest) for Radium-223 dichloride/placebo, Exemestane and Everolimus, respectively.
- Incidence rate of treatment-emergent AEs of concomitant medication of interest. [1]
- Incidence rate of treatment-emergent serious adverse events.
- Incidence rate of treatment-emergent drug-related serious adverse events , with respect to Radium-223 dichloride/placebo, Exemestane and Everolimus, respectively.
- The maximum severity of the TEAEs 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.
- 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.[2]
- Time to fracture is defined as the time (months) from the first dose date to the date of first fracture. Subjects without fractures are censored at the last visit date before or on the end of active follow-up period. Time to fracture will be summarized using Kaplan-Meier

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

estimates. Median survival time together with the 25th and 75th percentiles and associated 95% confidence intervals (CI) will be presented by treatment arm. Corresponding Kaplan-Meier curves will be generated by treatment arm.

- Competing risks for time to fracture is defined fracture as the primary event, death as competing risks. Patients who had fracture will be counted as primary event. Patients who died without fracture will be counted as competing risk.

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 period, all deaths during active follow up)

- Listing of subjects who died during treatment period: 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, 160 events are required to detect a 50% increase in SSE-FS. The expected study duration is 26.9 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 12.45 month long, enrollment for a total of 311 patients in the two treatment groups combined.

The power calculation does not take into account the administrative interim analysis, which was added in Protocol Amendment 8 (dated 18 MAY 2017). No alpha and no beta spending

adjustment is performed. There was no plan to stop the study for futility or superiority. However, an interim analysis may affect the operating characteristics.

The power calculation does not take into account the Sponsor decisions to discontinue enrollment in April 2018, and to curtail the study and conduct the primary analysis, independently of the number of SSE-FS achieved, with a cutoff date of 22JAN2020. Because the study is being ended and the primary analysis performed with less than the originally planned number of patients and events, the actual power of this study is expected to be less than planned.

6.7 Additional analyses planned to be reported outside the main report

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

6.8 Change of analysis[1]

After the unblinded administrative data review, conducted on 05 DEC 2017, the IDRC recommended discontinuing enrollment on 30 NOV 2017. Protocol Amendment 9, dated 03 APR 2018, allowed subjects who completed the EOT visit to be transitioned to the separate long-term follow-up study (BAY 88-8223 Study 16996 / NCT02312960), without further follow-up in the 17096 study.

The premature enrollment discontinuation, with subsequent reduced sample size, and the potentially reduced follow-up period limited the number of SSE-FS events observed. Therefore, the sponsor decided to conduct the primary analysis, independently of the number of SSE-FS achieved, with a cutoff date of 22JAN2020.

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 specified in previous sections are calculated for 95%% CI only. A hazard ratio with 2-sided 95% CI from a stratified Cox proportional hazards model will be performed as planned. The use of 80% CIs specified in the protocol will be limited to the primary endpoint main analysis.

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

- Age (≥55, <55)

[1] As of Amendment 4.

- 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
- Analysis of impact of baseline ideal body weight and total body weight on AEs

7. Document history and changes in the planned statistical analysis

- Signed SAP version 1.0 dated 25 October 2016, according to protocol version 1.0 up to amendment 5.0, 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 8, integrated protocol version 4.0, and dated 23 May 2017.

- Signed SAP version 3.0 dated 04 JAN 2019, according to protocol version 1.0 up to amendment 9, integrated protocol version 6.0, and dated 03 APR 2018.

- Signed SAP version 4.0 dated 14 FEB 2020, according to protocol version 1.0 up to amendment 10, integrated protocol version 7.0, and dated 04 DEC 2019.

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]. Baselga J, Campone M, Piccart M, Burris HA 3rd, Rugo HS, Sahmoud T, et al. Everolimus in postmenopausal hormone-receptor-positive advanced breast cancer. N Engl J Med. 2012;366(6):520-9. http://dx.doi.org/10.1056/NEJMoa1109653. Epub 2011 Dec 7.

[4]. Yardley et al. Everolimus plus exemestane in postmenopausal patients with HR+ breast cancer: BOLERO-2 final progression-free survival. AnalysisAdv Ther. 2013;30:870-84.

[5]. 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 9–1: 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

Table 9–1: 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)**
Not evaluable/Missing	Unequivocal progression	PD	Bone response date
Complete response	Not (all) evaluated/Missing	NE	
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)"

Title page

A phase II randomized, double-blind, placebo-controlled trial of radium-223 dichloride in combination with exemestane and everolimus versus placebo in combination with exemestane and everolimus when administered to metastatic HER2 negative hormone receptor positive breast cancer subjects with bone metastases

Short title: Study of radium-223 dichloride in combination with exemestane and everolimus versus placebo in combination with exemestane and everolimus in subjects with bone predominant HER2 negative hormone receptor positive metastatic breast cancer

Bayer study drug	BAY no. 88-8223 / Radium-223 dichloride / Xofigo		
[Study purpose:]	Efficacy and Safety		
Clinical study phase:	II	**Date:**	24 JUN 2020
Study No.:	BAY 88-8223 / 17096	**Version:**	1.0
Author:	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

Abbreviations

ALP	Alkaline Phosphate
CSR	Clinical Study Report
FAS	Full Analysis Set
rPFS	Radiological progression-free survival
SAP	Statistical Analysis Plan

1. Introduction

The statistical analysis plan (SAP), version 4.0, dated 14 FEB 2020, was considered the main SAP applied for analysis. Its analyses were pre-hoc, specified and applied prior to database lock of the study.

This Supplemental SAP version 1.0 describes analyses that were not included in the main SAP but may be used for clinical study report (CSR) or other submission documents. Changes to planned analyses in the main SAP are also described in this Supplemental SAP.

This Supplemental SAP version 1.0 is a supplement to SAP version 4.0 dated 14 FEB 2020.

Unless otherwise indicated in this supplemental SAP, the analyses described in this supplemental SAP should be considered post-hoc and may have been specified after database lock.

2. Study Objectives

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

The exploratory endpoints are:

- time to first on-study SSE;
- time to bone-ALP progression;
- bone-ALP response at Week 12 and 4 weeks (± 7 days) after last radium-223 dichloride/placebo dose;
- bone-specific rPFS;
- resource utilization;
- biomarker assessments,
- time to visceral metastases onset.

3. Study Design

Refer to the main SAP v4.0 dated 14 FEB 2020.

4. General Statistical Considerations

Refer to the main SAP v4.0 dated 14 FEB 2020.

4.4 Data Rules

Refer to the main SAP v4.0 dated 14 FEB 2020.

For the survival assessment-related main analysis of time-to-event endpoints which incorporate survival information, including primary endpoint SSE-FS, secondary endpoints OS and rPFS, and exploratory endpoint bone-specific rPFS, 17096 subjects who transferred to the 16996 study will have survival information from the 16996 study incorporated in determining the death date. Accordingly, subjects whose deaths were recorded in the 16996 study will have death events recorded at their respective death dates as documented in the 16996 study , In addition, for secondary endpoint OS, which incorporates last known alive date (LKAD) for censoring purposes, subjects who were still alive and followed up in 16996 as of the cutoff date will be censored at the last known alive date prior to or on the cutoff including survival assessments performed and documented in in 16996,

The above data rule for the above endpoints was implemented prior to database lock. Accordingly, it does not change the pre-hoc nature of the analyses of these endpoints as otherwise described in the main SAP.

Analyses of three supplemental exploratory endpoints defined in this supplemental SAP will also use data from 16996, including competing risks for time to radiological bone progression, competing risks for time to radiological non-bone progression, and summary of death events. For these endpoints. 17096 subjects who transferred to the 16996 study will have survival information from the 16996 study incorporated in determining the death date. Accordingly, subjects whose deaths were recorded in the 16996 study will have death events recorded at their respective death dates as documented in the 16996 study. These endpoints, and this data rule for them, are considered post hoc.

5. Analysis Sets

Refer to the main SAP v4.0 dated 14 FEB 2020.

6. Statistical Methodology

Refer to the main SAP v4.0 dated 14 FEB 2020.

6.1 Population characteristics

Refer to the main SAP v4.0 dated 14 FEB 2020.

6.2 Efficacy

Refer to the main SAP v4.0 dated 14 FEB 2020.

Subjects who completed the EOT visit were eligible to be transitioned to the separate long-term follow-up study (BAY 88-8223 Study 16996 / NCT02312960), without further follow-up in the 17096 study. For primary analyses (described in the main SAP v4.0 dated 14 FEB 2020) as well as post-hoc analyses described in the present SAP, the survival data, including death date and last-known alive date, obtained in 16996 will be extracted and merged to the 17096 data.

6.2.1 Primary efficacy variables

Refer to the main SAP v4.0 dated 14 FEB 2020.The primary analysis of SSE-FS, and any associated sensitivity, supplemental, and/or subgroup analyses specified in the main SAP, will be performed incorporating data from study 16996 as specified in Section 4.4 Data Rules. These analyses will otherwise be performed as specified in main SAP V4.0, Section 6.2.1. As the data rule was implemented prior to database lock, these analyses are considered pre-hoc.

In addition, a sensitivity analysis for SSE-FS will be performed using data from study 17096 only. All data from study 16996 will be excluded from this sensitivity analysis. This sensitivity analysis will otherwise be performed in the same manner as the primary analysis of SSE-FS as specified in main SAP V4.0, Section 6.2.1.2. This sensitivity analysis is considered post-hoc.

6.2.2 Secondary efficacy variables

Refer to the main SAP v4.0 dated 14 FEB 2020.

6.2.2.1 Definition

Refer to the main SAP v4.0 dated 14 FEB 2020.

6.2.2.2 Analysis of secondary efficacy endpoints

The primary analyses of OS and rPFS, and any associated sensitivity, supplemental, and/or subgroup analyses specified in the main SAP, will be performed incorporating data from study 16996 as specified in Section 4.4 Data Rules, These analyses will otherwise be performed as specified in main SAP V4.0, Section 6.2.1. As the data rule was implemented prior to database lock, these analyses are considered pre-hoc.

In addition, sensitivity analyses for OS and rPFS respectively will be performed using data from study 17096 only. All data from study 16996 will be excluded from these sensitivity analyses. These sensitivity analyses will otherwise be performed in the same manner as the primary analyses of OS and rPFS, respectively, as specified in main SAP V4.0, Section 6.2.1.2. These sensitivity analyses are considered post-hoc.

Time to pain progression was evaluated in subjects with baseline WPS ≤ 8. However, subjects with baseline WPS score >8 will also be included in the analysis by censoring at radium-223 dichloride or placebo treatment start date. Pain improvement rate was summarized with safety subjects with baseline WPS ≥ 2 as denominator. These analyses are considered pre-hoc, in which the analysis populations were not clearly specified in main SAP V4.0, but included in this supplement SAP.

6.2.3 Exploratory efficacy variables

6.2.3.1 Definition

Refer to the main SAP v4.0 dated 14 FEB 2020.

6.2.3.2 Analysis of exploratory efficacy endpoints defined in the main SAP

The primary analyses of bone-specific rPFS, and any associated sensitivity, supplemental, and/or subgroup analyses specified in the main SAP, will be performed incorporating data from study 16996 as specified in Section 4.4 Data Rules, These analyses will otherwise be performed as specified in main SAP V4.0, Section 6.2.1. As the data rule was implemented prior to database lock, these analyses are considered pre-hoc.

In addition, a sensitivity analysis for bone-specific rPFS respectively will be performed using data from study 17096 only. All data from study 16996 will be excluded from this sensitivity analysis. This sensitivity analysis will otherwise be performed in the same manner as the primary analyses of bone-specific rPFS, respectively, as specified in main SAP V4.0, Section 6.2.1.2. This sensitivity analysis is considered post-hoc.

For analysis of time to bone ALP progression, 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 18 weeks (2 × (8 weeks +1 week)) during the active follow-up period. The time window for active follow-up period is mis-specified in main SAP V4.0 section 6.2.3.1, but corrected in this supplement SAP.

6.2.3.3 Definition of additional exploratory efficacy variables

The additional exploratory efficacy variables are specified below:

- Time to radiological bone progression (Kaplan-Meier)
- Time to radiological non-bone progression (Kaplan-Meier).
- Competing risks for time to radiological bone progression (death and radiological non-bone progression as competing risks)
- Competing risks for time to radiological non-bone progression (death and radiological bone progression as competing risks)
- Summary of death events with radiographic bone and non-bone progression as competing risks
- Summary of visceral disease

Derivation of three out of the above six variables, including competing risks for time to radiological bone progression, competing risks for time to radiological non-bone progression, and summary of death events, will also include the addition of 16996 data as described in Section 4.4 Data Rules.

Time to Radiological Bone Progression (Kaplan-Meier) is defined as the time (months) from randomization date to the date of radiological bone progression (according to the

mRECIST1.1 criteria), as documented by the investigator. Radiological bone progression occurs when progression by bone scan is detected.

Time to Radiological Bone progression = (End Date – Randomization Date +1)/30.44

Table 6–1: Time to Radiological Bone progression censoring rules (Kaplan-Meier analysis)

Situation	End Date	Censored	Reason for Censoring
No baseline or post-baseline radiological bone assessment (bone scan)	Randomization Date	Yes	No baseline or post-baseline bone radiological assessment.
Superscan at baseline	Randomization Date	Yes	Superscan at baseline
Subject had a radiological bone assessment of PD (no two consecutive missed radiological bone assessments)	Date of first bone PD	No	N/A
PD immediately after two or more consecutive missed radiological bone assessments	Date of last radiological bone assessment before missed assessments	Yes	Missed two or more consecutive bone scan assessments
Subject discontinuation from study for other than PD	Last radiological bone assessment date	Yes	Subject discontinued from study due to a reason other than bone PD
Death during the study before first radiological bone PD assessment	Randomization Date	Yes	No radiological bone progression observed
Subjects discontinued from study due to PD, but no documented date of bone PD	Date of last radiological bone assessment	Yes	Subjects discontinued from study due to PD, but no documented date of bone PD
Subjects still on study at the time of data cutoff without bone PD	Last radiological bone assessment before data cutoff	Yes	Subject is still alive without bone PD
New systemic anticancer treatment started	Date of last radiological bone assessment before starting new systemic anticancer treatment	Yes	New systemic anticancer treatment started

Time to Radiological non-Bone Progression (Kaplan-Meier) is defined as the time (months) from randomization date to the date of radiological non-bone (soft-tissue or visceral lesions) progression (according to the mRECIST1.1 criteria), as documented by the investigator.

Radiological non-bone progression occurs when progression by non-bone scan is detected.

Time to Radiological non-Bone progression = (End Date – Randomization Date +1)/30.44

Table 6–2: Time to Radiological non-Bone progression censoring rule (Kaplan-Meier analysis)

Situation	End Date	Censored	Reason for Censoring
No baseline or post-baseline radiological non-bone assessment (non-bone scan)	Randomization Date	Yes	No baseline or post-baseline non-bone radiological assessment.
Subject had a radiological non-bone assessment of PD (no two consecutive missed radiological non-bone assessments)	Date of first non-bone PD	No	N/A
Non-bone PD immediately after two or more consecutive missed radiological non-bone assessments	Date of last radiological non-bone assessment before missed assessments	Yes	Missed two or more consecutive non-bone scan assessments
Subject discontinuation from study for other than non-bone PD	Last radiological bone assessment date	Yes	Subject discontinued from study due to a reason other than non-bone PD
Death during the study before first radiological non-bone PD assessment	Randomization date	Yes	No radiological non-bone progression observed
Subjects discontinued from study due to non-bone PD, but no documented date of non-bone PD	Date of last radiological non-bone assessment	Yes	Subjects discontinued from study due to PD, but no documented date of non-bone PD

Table 6–2: Time to Radiological non-Bone progression censoring rule (Kaplan-Meier analysis)

Situation	End Date	Censored	Reason for Censoring
Subjects still on study at the time of data cutoff without non-bone PD	Last radiological non-bone assessment before data cutoff	Yes	Subject is still alive without non-bone PD
New systemic anticancer treatment started	Date of last radiological non-bone assessment before starting new systemic anticancer treatment	Yes	New systemic anticancer treatment started

Competing risks for time to radiological bone progression and competing risks for time to radiological non-bone progression

Once time to radiological bone progression and time to radiological non-bone progression event and Kaplan-Meier censoring rules are defined above.

Competing risks for time to radiological bone progression is defined radiological bone progression as the primary event, death and radiological non-bone progression as two competing risks. Subjects who experienced radiological bone progression before or on the first occurrence of radiological non-bone progression or death will be counted as primary event. Subjects who died or had radiological non-bone progression without radiological bone progression will be counted as competing risks. Subjects with neither a primary nor a competing risk event will be censored. Censoring rules, including rules addressing missing assessments, are described below.

Competing risks for time to radiological non-bone progression is defined radiological non-bone progression as the primary event, death and radiological bone progression as two competing risks. Subjects who experienced radiological non-bone progression before or on the first occurrence of radiological bone progression or death will be counted as primary event. Subjects who died or had radiological bone progression without radiological bone progression will be counted as competing risks. Subjects with neither a primary nor a competing risk event will be censored. Censoring rules, including rules addressing missing assessments, are described below.

The competing risks are defined with combination the outcome of different type of events/censoring. The outcome combination

logic is as below:

(1) Rules for addressing subsequent therapy and missing assessments (missing assessment windowing):

- If there was a new anti-cancer therapy, then all tumor assessments and any death event following the date of the first anti-cancer therapy (both bone and non-bone) are discarded.
- For each type of tumor assessment (bone and non-bone), considered separately, if there are two consecutive missing assessments (gap between 2 assessments of > 18 weeks) for that assessment type prior to the PD (PD is > 18 weeks after last non-PD assessment) for the type, then the PD is discarded.
- For each type of tumor assessment (bone and non-bone), considered separately, if there are two consecutive missing assessments (gap of > 18 weeks) for that assessment type prior to a death event with no prior progression of that type, then the death event is discarded for purposes of evaluating the progression type (Note: the death event may count for the other progression type if assessments did not result in a gap).
- After applying the above rules, if a PD event occurs >18 weeks after the last-dated assessment of the other tumor type, and the other tumor type has no PD, then the PD is discarded.

(2) Rules for addressing simultaneity: . After applying the above rules, if the last-dated remaining bone scan is within ±7 days of the last-dated remaining non-bone tumor assessment (either assessment occurred 7 or fewer days after the other), then the two assessments will be regarded as effectively occurring simultaneously. This simultaneity windowing concept addresses the potential for tumor assessments of different types to occur on different dates even though they are effectively part of the same tumor assessment visit, and to reduce the possibility of minor scheduling anomalies biasing the outcomes. To implement this, the following table (Table 6–3) determines the outcome and date of each endpoint:

Table 6–3: Rules for addressing simultaneity

Endpoint	**Earlier of Last Assessments (Bone or non-Bone)**	**Earlier Assessment Outcome**	**Later Assessment Outcome**	**Outcome**	**Date**
TT Bone Progression	Bone Assessment	PD	PD	Event	Bone Assessment
	Bone Assessment	PD	Non-PD	Event	Bone Assessment
	Bone Assessment	Non-PD	PD	Competing Risk Event	Non-Bone Assessment
	Bone Assessment	Non-PD	Non-PD	Censor	Bone Assessment

Table 6–3: Rules for addressing simultaneity

Endpoint	Earlier of Last Assessments (Bone or non-Bone)	Earlier Assessment Outcome	Later Assessment Outcome	Outcome	Date
	Non-Bone Assessment	PD	PD	Event	Bone Assessment
	Non-Bone Assessment	PD	Non-PD	Competing Risk Event	Non-Bone Assessment
	Non-Bone Assessment	Non-PD	PD	Event	Bone Assessment
	Non-Bone Assessment	Non-PD	Non-PD	Censor	Bone Assessment
TT Non-Bone Progression	Bone Assessment	PD	PD	Event	Non-Bone Assessment
	Bone Assessment	PD	Non-PD	Competing Risk Event	Bone Assessment
	Bone Assessment	Non-PD	PD	Event	Non-Bone Assessment
	Bone Assessment	Non-PD	Non-PD	Censor	Non-Bone Assessment
	Non-Bone Assessment	PD	PD	Event	Non-Bone Assessment
	Non-Bone Assessment	PD	Non-PD	Event	Non-Bone Assessment
	Non-Bone Assessment	Non-PD	PD	Competing Risk Event	Bone Assessment
	Non-Bone Assessment	Non-PD	Non-PD	Censor	Non-Bone Assessment

(3) Rules for addressing precedence (after applying both windowing rules above): Rules are based on the following table (Same for time to bone progression/bone assessments and time to non-bone progression/non-bone assessments respectively). The window refered to in the table below (i.e., Table 6–4) is a ±7 day window:

Table 6–4: Rules for addressing precedence

Last Endpoint	Last Other	Other	Death	Outcome	Date

(Bone or Non-Bone) Assessment	**(Non-Bone or Bone) Assessment**	**Precedence**			
PD	PD	Before, outside window	After or none	Competing Risk Event (Other)	Other Assessment
Any	Any	Window	After or none	See Table 6–3	3
PD	PD	After, outside window	After or none	Event	Endpoint Assessment
PD	Non-PD	After, outside window	After or none	Event	Endpoint Assessment
PD	Non-PD	Before, outside window	After or none	Event	Endpoint Assessment
Non-PD	PD	Before, Outside Window	After or none	Competing Risk Event (Other)	Other Assessment
Non-PD	PD	After, Outside window	After or none	Censored	Endpoint Assessment
Non-PD	Non-PD	Any, outside Window	After	Competing Risk Event (Death)	Death
Non-PD	Non-PD	Any, outside Window	none	Censored	Endpoint Assessment

6.2.3.4 Analysis of additional exploratory efficacy endpoints

For Kaplan-Meier time to radiological bone progression and time to radiological none-bone progression, analysis will be performed similarly to the primary endpoint. To compare the two treatment groups, a stratified log-rank test will be performed, stratified by the same stratification factors as randomization. The hazard ratio (radium-223 dichloride / placebo) and its 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 95% 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.

For competing risks variables, competing-risk time to radiological bone progression and competing-risk time to radiological none-bone progression, the hazard ratio (radium-223 dichloride / placebo) for and its 95% CIs for time to event with competing risks will be estimated from the subdistribution proportional hazard model, stratified by the same factors as for SSE-FS analysis. Time to event with competing risks efficacy analysis will be performed to compare two treatment groups using Gray's test[1] stratified by the same stratification factors as randomization. The estimates of the nonparametric cumulative incidence functions

will be presented for each treatment group: N, total censored, total event, total competing risks, rate of subjects who developed an event at pre-specified months. The cumulative incidence function curves will be generated for each treatment group.

Due to the small number of death events with radiological bone and non-bone progression ask competing risks, summary of death events with radiographic bone and non-bone progression as competing risks will be provided. No time to event analysis will be performed.

The number of subjects (N and percenatages) with visceral disease will be summarized with bone rPFS, death and non-bone rPFS events, respectively.

6.3 Safety

Refer to the main SAP v4.0 dated 14 FEB 2020.

6.4 Pharmacokinetics

Refer to the main SAP v4.0 dated 14 FEB 2020.

6.5 Biomarker analyses

Refer to the main SAP v4.0 dated 14 FEB 2020.

6.6 Sample size estimation

Refer to the main SAP v4.0 dated 14 FEB 2020.

6.7 Additional analyses planned to be reported outside the main report

Refer to the main SAP v4.0 dated 14 FEB 2020.

6.8 Change of analysis

Refer to the main SAP v4.0 dated 14 FEB 2020.

Additional change of analysis:

Due to the limitation of data collection, overall duration of single standard of care hormonal treatment in follow up period and dosage of hormonal treatment, which were pre-hoc analyses specified in main SAP V4.0, will not be summarized in CSR.

7. Document history and changes in the planned statistical analysis

Refer to the main SAP v4.0 dated 14 FEB 2020.

8. References

Refer to the main SAP v4.0 dated 14 FEB 2020.

[1] Gray, R. J. (1988). “A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk.” *Annals of Statistics* 16:1141–1154.

9. Appendices

Refer to the main SAP v4.0 dated 14 FEB 2020.