

---

## ***STATISTICAL ANALYSIS PLAN***

---

**A Randomized, Double-Blind, Placebo-Controlled, Multi-Center Phase 3 Study of Denosumab as Adjuvant Treatment for Women with Early-Stage Breast Cancer at High Risk of Recurrence (D-CARE)**

**Protocol Number 20060359**

**Version:** 3.0  
**Date:** 15 November, 2017  
**Authors:** PPD [REDACTED], PPD [REDACTED]

NCT Number: 1077154  
This NCT number has been applied to the document  
for purposes of posting on clinicaltrials.gov

## Table of Contents

Table of Abbreviations .....	6
1. Introduction .....	8
2. Objectives .....	8
2.1 Primary Objective .....	8
2.2 Secondary Objectives.....	8
2.3 Exploratory Objectives.....	8
3. Study Overview .....	9
3.1 Study Design .....	9
3.2 Sample Size .....	11
4. Study Endpoints and Covariates.....	12
4.1 Primary and Secondary Efficacy Endpoints .....	12
4.2 Safety Endpoints .....	13
4.2.1 Adverse Events (AEs) .....	13
4.2.2 Laboratory Values .....	14
4.2.2.1 Serum Chemistry .....	14
4.2.2.2 Hematology.....	14
4.2.3 Anti-denosumab Antibody .....	14
4.2.4 Vital Signs.....	14
4.2.5 Performance Status.....	15
4.3 Exploratory Endpoints .....	15
4.3.1 Efficacy Exploratory Endpoints.....	15
4.3.2 Patient Reported Outcome (PRO) Endpoints.....	16
4.3.3 Denosumab Serum Concentration Levels.....	16
4.3.4 Bone Turnover Markers.....	17
4.4 Covariates .....	17
5. Definitions .....	18
5.1 Basic Definition.....	18
5.2 Study Points of Reference.....	18
5.3 Study Dates .....	18
5.4 Study Time Intervals.....	19
5.5 Subject Dispositions .....	19
5.6 Study-specific Definitions .....	20
5.7 Derived Variables .....	25
6. Analysis Data Sets .....	31
6.1 Data Subsets .....	31
6.1.1 Full Analysis Set (FAS).....	31
6.1.2 Per Protocol Set .....	31
6.1.3 Safety Analysis Set.....	33

6.1.4	Pharmacokinetic (PK) Analysis Set	33
6.2	Subset for Interim Analysis	33
6.3	Subgroup Analyses	34
7.	Interim Analysis and Early Stopping Guidelines	34
8.	Data Screening and Acceptance	37
8.1	General Principles	37
8.2	Data Handling and Electronic Transfer of Data	37
8.3	Handling of Missing and Incomplete Data	37
8.3.1	Missing Dates	37
8.3.2	Assayed Values Below or Above Quantifiable Limits of Bone Turnover Marker (BTM)	39
8.3.3	Missing / Unknown Values of Covariates	39
8.3.4	Missing PRO data	39
8.3.5	Missing Pathological Complete Response	39
8.4	Assessment of Outliers	39
8.5	Distributional Characteristics	40
8.6	Validation of Statistical Analysis	40
9.	Statistical Methods of Analysis	40
9.1	General Principles	40
9.2	Subject Accountability	41
9.3	Demographic and Baseline Characteristics	41
9.3.1	Demographics	42
9.3.2	Baseline Characteristics	42
9.3.3	Bone-specific Medication History	43
9.3.4	Fracture History	43
9.3.5	Baseline PRO Characteristics and Analgesic Use	43
9.4	Safety Analyses	44
9.4.1	Adverse Events (AEs)	44
9.4.1.1	Events of Interests (EOIs)	44
9.4.1.2	Osteonecrosis of the Jaw	45
9.4.1.3	Atypical Femoral Fracture	45
9.4.2	Investigational Product Exposure	45
9.4.3	Concomitant Medications and Anti-neoplastic Therapies/Surgeries	46
9.4.4	ECOG Performance Status	46
9.4.5	Vital Signs	46
9.4.6	Immunogenic Response	46
9.4.7	Clinical Laboratory Results	46
9.5	Efficacy Analyses	47
9.5.1	Primary and Secondary Efficacy Endpoints	48
9.5.1.1	Estimation of Outcomes	48

9.5.1.2	Hypotheses to Be Tested.....	48
9.5.1.3	Primary Analysis .....	49
9.5.1.4	Supportive Analyses .....	50
9.5.1.5	Subgroup Analyses.....	51
9.5.1.6	Analysis of Covariates .....	52
9.5.2	Exploratory Endpoints .....	52
9.5.2.1	Time to Event Exploratory Endpoints.....	52
9.5.2.2	Analgesic Score .....	53
9.5.2.3	PRO Analysis.....	53
9.5.2.4	Breast Density.....	55
9.5.2.5	Pathological Complete Response.....	55
9.5.2.6	Residual Invasive Tumor Size .....	55
9.5.2.7	Serum Denosumab Trough Concentrations.....	55
9.5.2.8	Blood Biomarkers and Bone Turnover Markers.....	55
10.	Final Analysis of Data From Ltfu .....	56
11.	Literature Citations / References.....	57
12.	Appendices .....	59
12.1	Technical Detail/Supplemental Information Regarding Statistical Procedures and Programs.....	59
12.1.1	Dates .....	59
12.1.2	Visit Windows .....	59
12.2	Reference Values / Toxicity Grades .....	63
12.3	Patient Report Outcome Forms/Instruments .....	64
12.3.1	Patient Report Outcome Forms.....	64
12.3.2	Patient Report Outcome Instruments .....	64
12.3.2.1	Reliability and Validity of PRO Instruments .....	64
12.3.2.2	PRO Question Mapping .....	64
12.3.2.3	Scoring Algorithm for PROs.....	65

**List of Tables**

Table 1.	Primary and Secondary Efficacy Endpoints .....	13
Table 2.	Adverse Event.....	13
Table 3.	Serum Chemistry .....	14
Table 4.	Hematology .....	14
Table 5.	Anti-denosumab Antibody .....	14
Table 6.	Vital Signs .....	14
Table 7.	Performance Status .....	15
Table 8.	Efficacy Exploratory Endpoints .....	15

---

Table 9. PRO Endpoints .....	16
Table 10. Bone Turnover Markers .....	17
Table 11. Decision Rule (1-sided) at Interim/Primary Analyses .....	36
Table 12. Imputation Rules for Incomplete Dates .....	38

**Table of Abbreviations**

Abbreviation	Definition
AE	adverse events
AFF	atypical femoral fracture
AJCC	American Joint Committee on Cancer
$\alpha$ CTx	C-telopeptide of type I collagen
AQL	above quantifiable level
BMFS	Bone metastasis-free survival
BPI-SF	Brief Pain Inventory – Short Form
BQL	below quantifiable level
BI-RADS	Breast Imaging-Reporting and Data System
BSAP	bone specific alkaline phosphatase
BTM	bone turnover marker
CDM	clinical data management
CI	confidence interval
CRF	case report form
CRO	clinical research organization
CT	computed tomography
CTCAE	common terminology criteria for adverse events
DBP	diastolic blood pressure
DMC	data monitoring committee
DFS	disease-free survival
DRFS	distant recurrence-free survival
ECOG	Eastern Cooperative Oncology Group
EDC	Electronic Data Capture
EOD	extraosseous disease
EOI	event of interest
EOS	end of study
EQ-5D	EuroQol –5 Dimensions
FAS	full analysis set
GEE	generalized estimating equation
HLT	high level term
IP	investigational product
LTFU	Long-term follow-up
MRI	magnetic resonance imaging
OME	morphine equivalent

Abbreviation	Definition
ONJ	osteonecrosis of the jaw
OS	overall survival
PRO	patient reported outcomes
RITS	residual invasive tumor size
SBP	systolic blood pressure
SPECT	single-photon emission CT
SRE(s)	skeletal-related event(s)
uNTx	Urinary N-telopeptide of type I collagen
VAS	visual analogue scale

## 1. Introduction

This document outlines the planned statistical analyses for data collected within the scope of Amgen denosumab protocol 20060359, entitled “A Randomized, Double-Blind, Placebo-Controlled, Multi-Center Phase 3 Study of Denosumab as Adjuvant Treatment for Women with Early-Stage Breast Cancer at High Risk of Recurrence (D-CARE)”. This statistical analysis plan (SAP) applies to the protocol amendment 4 dated on October 17, 2016. The purpose of this plan is to provide general, and in some instances, specific guidelines from which the analysis will proceed.

The primary analysis will include efficacy and safety data up to the primary analysis data cut-off date. The analysis time period/point for each endpoint is specified in [Section 4](#). Final analysis will be conducted at the end of the long-term follow-up (LTFU) phase ([Section 10](#)).

## 2. Objectives

### 2.1 Primary Objective

To compare the treatment effect of denosumab with that of placebo on prolonging bone metastasis-free survival (BMFS) in subjects with early-stage breast cancer at high risk of disease recurrence.

### 2.2 Secondary Objectives

To compare the treatment effect of denosumab with that of placebo on:

- Disease-free survival (DFS) in the full study population
- DFS in the postmenopausal subset
- Overall survival (OS)
- Distant recurrence-free survival (DRFS)

### Safety Objectives

- To assess the safety and tolerability of denosumab compared with placebo

### 2.3 Exploratory Objectives

To evaluate the treatment effect of denosumab compared with placebo on:

- Time to first bone metastasis (excluding deaths)
- Time to bone metastasis as site of first recurrence
- Time to disease recurrence
- Time to distant recurrence
- Time to first on-study fracture (vertebral or non-vertebral fracture)
- Time to first on-study skeletal-related event (SRE; following the development of bone metastasis)



- Time to first on-study SRE or hypercalcemia (following the development of bone metastasis)
- Time to first on-study symptomatic bone metastasis
- Brief Pain Inventory - Short Form (BPI-SF) 'worst' pain score
- BPI-SF pain severity and pain interference scales
- EQ-5D health index and visual analogue scale (VAS) scores
- Analgesic use
- Breast density
- Pathological variables in tumor tissue from neoadjuvant subjects after surgery
- Pharmacokinetics
  - Serum trough levels of denosumab
- Pharmacodynamic response
  - Levels and dynamics of bone turnover markers and potential other pharmacodynamic markers

To investigate other potential biomarker development (eg, biochemical levels and abundance of drug targets) by biochemical analysis of blood and/or tumor tissue, and correlate with treatment outcomes

To investigate the association of tumor genetic variations in cancer genes, drug/pathway target genes, and/or other biomarker genes with treatment outcomes

To investigate the genetic variation in cancer genes, drug/pathway target genes, and/or other biomarker genes, and correlate with treatment outcomes (optional; requires separate informed consent)

### **3. Study Overview**

#### **3.1 Study Design**

This is an international, phase 3, randomized, double-blind, placebo-controlled study in women with early-stage (stage II or III) breast cancer at high risk of disease recurrence. Approximately 4,500 subjects will be randomized in a 1:1 ratio to receive denosumab 120 mg or matching placebo. Randomization will be stratified based on the following criteria at the time of study entry:

- Breast cancer therapy / Lymph node (LN) status: neo-adjuvant therapy / any LN status versus [vs] adjuvant therapy / LN negative (based on axillary LN dissection, or based on sentinel node [SN] status) vs adjuvant therapy / LN positive
- Hormone receptor (estrogen receptor [ER] / progesterone receptor [PR]) status: ER and/or PR positive vs ER and PR negative
- Human epidermal growth factor receptor 2 (HER-2) status: HER-2 positive vs HER-2 negative

- Age: < 50 vs  $\geq$  50
- Region: Japan vs Other

Subjects will receive denosumab 120 mg or matching placebo subcutaneously (SC) every 4 weeks (Q4W;  $\pm$  7 days) for approximately 6 months followed by denosumab 120 mg or matching placebo SC every 3 months (Q3M; ie, every 12 weeks  $\pm$  14 days) for 4½ years (approximately 54 months), for a total treatment duration of 5 years (approximately 60 months). As there are approved bone-targeted therapies for the treatment of bone metastasis, subjects who develop bone metastasis will permanently discontinue investigational treatment, complete the End of Treatment Phase (EOTP) visit, and enter LTFU upon documented evidence of bone metastasis (per central imaging analysis or biopsy). For an individual subject, the maximum duration of LTFU after completing the EOTP visit is 5 years (total study duration for an individual subject is up to 10 years from the date of randomization). During the first approximate 6 months of treatment, investigational product (denosumab or matching placebo) may be administered Q3W to subjects receiving or scheduled to receive Q3W adjuvant/neoadjuvant chemotherapy, to minimize patient clinic visits and enhance adherence to treatment. There must be at least a 3-week interval between administrations of investigational product. All subjects will be required to receive vitamin D and calcium supplementation at standard doses (at least 500 mg calcium and at least 400 IU of vitamin D), unless documented hypercalcemia develops on study.

After completing the treatment phase of the study, subjects will then enter LTFU for approximately 5 years (60 months) following the EOTP visit. The study requirements during LTFU are different before and after primary analysis data cut-off date. Specifically, before the primary analysis data cut-off date, all subjects (including subjects with documented evidence of bone metastasis) should be followed by clinic visit or telephone contact every 3 months (every 12 weeks  $\pm$  14 days) for approximately 6 months (ie, first 2 visits of LTFU following EOTP) to assess clinical fractures/SREs, hypercalcemia of malignancy, breast cancer therapy, concomitant medications, patient-reported outcomes [BPI-SF, EQ-5D], disease recurrence status (until documented evidence of bone metastasis [per central imaging analysis or biopsy], not required for subjects with documented evidence of bone metastasis), and survival. All subjects without documented evidence of bone metastasis following the completion of the initial 6 month LTFU period should be followed by clinic visit or telephone contact every 3 months (every 12 weeks  $\pm$  14 days) to assess clinical fractures/SREs,

hypercalcemia, disease recurrence status (until documented evidence of bone metastasis [per central imaging analysis or biopsy]), and survival. Subjects with documented evidence of bone metastasis including subjects whose bone metastasis is confirmed after the initial 6 month (per central imaging analysis or biopsy): following the completion of the initial 6 month LTFU period should be followed by clinic visit or telephone contact every 6 months ( $\pm$  1 month) to assess SREs, hypercalcemia, and survival. Following the primary analysis data cut-off date, all subjects should complete the EOTP visit and will be followed by clinic visit or telephone contact every 6 months ( $\pm$  1 month) for survival only.

A data monitoring committee (DMC) external to Amgen and external to Daiichi Sankyo Co., Ltd. will review safety data at regular intervals, as outlined in the DMC charter (approximately twice yearly; the start date will depend on subject accrual rates). This external DMC will conduct 1 interim analysis of efficacy after approximately 367 subjects have developed bone metastasis or died, and approximately 555 subjects have developed any disease recurrence or died, whichever comes last. To ensure timely conduct of the interim analysis, the interim analysis may be conducted early, as soon as approximately 6 months after the sufficient event number has been reached for either BMFS or DFS in the full study population, even if the event number for the other endpoint has not been reached.

The primary analysis will be conducted after all enrolled subjects have had the opportunity to complete 5 years of treatment from study day 1.

### **3.2 Sample Size**

The sample size calculation is based on the primary endpoint, BMFS, and is estimated using EAST 6 or above. Estimates of the event rates for BMFS and DFS for the target population are based on (Colleoni et al., 2000) and on recent published and unpublished data, including (Francis et al., 2008, Martin et al., 2005) for the HER-2 negative population, (Romond et al., 2005, Slamon et al, 2006) for the HER-2 positive population, (Howell et al., 2005, Thurlimann et al., 2005) for the population receiving adjuvant hormonal therapy, and (Coleman et al., 2014) for the use of adjuvant zoledronic acid in stage II-III breast cancer. Based on the reported event rates, the event rate proportions for disease recurrence in the bone, the viscera, or death reported by (Colleoni et al., 2000) and (Coleman et al. 2014), and considering the availability of some breast cancer adjuvant therapies, the BMFS rates at three years is expected to be approximately 90.5% in the control group. Disease recurrence risk after primary breast

cancer treatment can vary over time ([Cheng et al, 2012](#)). For DFS, it is assumed that the time to event (death or disease recurrence) in the control group follows a piece-wise exponential distribution with a risk of about 5.5% per year for the first 3 years and 3.2% per year afterwards in the full study population, and a piece-wise exponential distribution with a risk of about 5.6% per year for the first 3 years and 3.7% per year afterwards in the postmenopausal subset. It is anticipated that the true hazard ratio of denosumab compared with placebo is 0.83 in the full study population and 0.76 in the postmenopausal subset.

The primary analysis will be conducted after all enrolled subjects have had the opportunity to complete 5 years of treatment from study day 1. With 4,500 subjects, an enrollment period of approximately 27 months, the study is estimated to reach the primary analysis data cut-off date in approximately 87 months (ie, 7 years and 3 months) from the start of enrollment.

With a loss to follow-up rate of 6% per year, for BMFS, if the true hazard ratio is 0.8, the power to detect superiority of denosumab over placebo in the primary analysis is estimated to be approximately 80%.

#### **4. Study Endpoints and Covariates**

##### **4.1 Primary and Secondary Efficacy Endpoints**

The primary and secondary efficacy endpoints are listed in [Table 1](#).

**Table 1. Primary and Secondary Efficacy Endpoints**

	Endpoints	Analysis Time Period
Primary	Bone metastasis-free survival (determined by the time from randomization to the first observation of bone metastasis or death from any cause)	Randomization through the primary analysis data cut-off date
Secondary	<ul style="list-style-type: none"> <li>• Disease-free survival in the full study population (determined by the time from randomization to the first observation of disease recurrence or death from any cause)</li> <li>• Disease-free survival in the postmenopausal subset (determined by the time from randomization to the first observation of disease recurrence or death from any cause)</li> <li>• Overall survival (determined by the time from randomization to death from any cause)</li> <li>• Distant recurrence-free survival (determined by the time from randomization to the first observation of distant metastasis or death from any cause)</li> </ul>	

**4.2 Safety Endpoints**

**4.2.1 Adverse Events (AEs)**

**Table 2. Adverse Event**

Parameter	Analysis Time Period
Treatment-emergent adverse events:  All AEs AEs leading to investigational product (IP) discontinuation AEs leading to study discontinuation Serious AEs Serious AEs leading to IP discontinuation Serious AEs leading to study discontinuation CTCAE Grade 3, 4, and 5 AEs Fatal AEs All AEs related to IP Adverse events of interest (EOIs)	Study day 1 through 30 days after the last dose for subjects who discontinued/completed IP or the primary analysis data cut-off date, whichever is earlier.

Note: The Common Terminology Criteria for Adverse Events (CTCAE Version 3.0 or higher) will be used.

## 4.2.2 Laboratory Values

### 4.2.2.1 Serum Chemistry

**Table 3. Serum Chemistry**

Parameter	Analysis Time Point
Serum chemistry (Recorded values and changes from baseline): Calcium / Albumin-adjusted calcium (also percent change) Serum 25 (OH) Vitamin D (baseline only) Magnesium Phosphorus (also percent change)	Baseline through the primary analysis data cut-off date by visit

### 4.2.2.2 Hematology

**Table 4. Hematology**

Parameter	Analysis Time Point
Hematology (Recorded values): Hemoglobin Platelets White blood cells (WBC) Neutrophils	Baseline

### 4.2.3 Anti-denosumab Antibody

**Table 5. Anti-denosumab Antibody**

Parameter	Analysis Time Point
Anti-denosumab Antibody (binding/neutralizing)	Study day 1 through the primary analysis data cut-off date

### 4.2.4 Vital Signs

**Table 6. Vital Signs**

Parameter	Analysis Time Point
Vital signs (Recorded values): Systolic blood pressure (SBP) Diastolic blood pressure (DBP) Pulse Respiration rate Temperature	Baseline through the primary analysis data cut-off date by visit

#### 4.2.5 Performance Status

**Table 7. Performance Status**

Parameter	Analysis Time Point
Performance status (Recorded values and change from baseline): Eastern Cooperative Oncology Group (ECOG) performance scale	Baseline through the primary analysis data cut-off date by visit

#### 4.3 Exploratory Endpoints

##### 4.3.1 Efficacy Exploratory Endpoints

**Table 8. Efficacy Exploratory Endpoints**

Parameter	Analysis Time Period
<ul style="list-style-type: none"> <li>Time to first bone metastasis (excluding deaths) (determined by the time from randomization to the first observation of bone metastasis as site of first or subsequent disease recurrence)</li> <li>Time to bone metastasis as site of first recurrence (determined by the time from randomization to the first observation of bone metastasis as site of first disease recurrence; excluding death)</li> <li>Time to disease recurrence (determined by the time from randomization to the first observation of disease recurrence; excluding death)</li> <li>Time to distant recurrence (determined by the time from randomization to the first observation of distant metastasis; excluding death)</li> <li>Time to first on-study fracture (determined by the time from randomization to the first observation of vertebral or non-vertebral fracture prior to the development of bone metastasis)</li> <li>Time to first on-study SRE (determined by the time from randomization to the first observation of SRE following the development of bone metastasis)</li> <li>Time to first on-study SRE or hypercalcemia (determined by the time from randomization to the first observation of SRE or hypercalcemia following the development of bone metastasis)</li> <li>Time to first on-study symptomatic bone metastasis (determined by the time from randomization to the first observation of bone metastasis which is accompanied by symptom at time of detection)</li> <li>Breast density (recorded value, and change from baseline)</li> <li>Pathological complete response for neo-adjuvant subjects</li> <li>Residual invasive tumor size for neo-adjuvant subjects</li> </ul>	Randomization through the primary analysis data cut-off date

### 4.3.2 Patient Reported Outcome (PRO) Endpoints

Recorded values and changes from baseline will be provided using descriptive statistics for each PRO parameter.

**Table 9. PRO Endpoints**

Parameter	Analysis Time Point
<b>Analgesic Use</b> Analgesic score Proportion of subjects with analgesic score $\geq 3$ (in subjects with baseline scores 0-2)	Baseline through the primary analysis data cut-off date (details in <a href="#">Section 9.5.2.2</a> )
<b>BPI-SF</b> “Worst” pain score Pain severity score Pain interference score Time to $\geq 2$ -point increase from baseline in “Worst” pain score (in subjects with baseline scores 0-8) Time to $> 4$ -point in “Worst” pain score (in subjects with baseline scores 0-4) Proportion of subjects with $\geq 2$ -point increase from baseline in “Worst” pain score (in subjects with baseline scores 0-8) Proportion of subjects with “Worst” pain score $> 4$ -point (in subjects with baseline scores 0-4)	Baseline through the primary analysis data cut-off date (details in <a href="#">Section 9.5.2.3</a> )
<b>EQ-5D</b> Health Index score Visual Analogue Scale (VAS)	Baseline through the primary analysis data cut-off date (details in <a href="#">Section 9.5.2.3</a> )

### 4.3.3 Denosumab Serum Concentration Levels

Serum denosumab concentration levels will be obtained from a subset of subjects prior to administration of the investigational products and on study. For the primary analyses, serum denosumab concentration levels will be summarized by visit up to the primary analysis data cut-off date.



#### 4.3.4 Bone Turnover Markers

**Table 10. Bone Turnover Markers**

Parameter	Analysis Time Point
BTMs (Recorded values, change and percent change from baseline):	Study day 1 and then by visit
Urinary N-telopeptide corrected by urine creatinine (uNTx/Cr)	
Urinary C-telopeptide corrected by urine creatinine (αCTx/Cr)	
Bone specific alkaline phosphatase (BSAP)	

#### 4.4 Covariates

The relationship of the following covariates to the primary and key secondary efficacy endpoints may be explored:

- Breast cancer therapy / Lymph node status (neo-adjuvant therapy / any LN status, adjuvant therapy / LN negative, and adjuvant therapy / LN positive)
- Hormone receptor (ER/PR) status (positive, negative)
- HER-2 status (positive, negative)
- Age (continuous)
- Menopausal status (premenopausal, postmenopausal)
- Postmenopausal status (>5 years since menopause, other [including premenopause, ≤5 years since menopause, and unknown])
- Breast cancer AJCC stage (IA or IB or IIA, IIB, IIIA, IIIB and above)
- Primary tumor size (T stage: T0 or T1 or T1a -T1c or T1mi or Tis, T2, T3, T4 or T4a-T4d)
- Lymph node status (N-stage: N0 or pN0 or pN0 (i-) or pN0 (i+) or pN0 (mol-) or pN0 (mol+) or pN1mi, N1 or pN1 or pN1a - pN1c, N2 or N2a or N2b or pN2 or pN2a or pN2b, N3 or N3a - N3c or pN3 or pN3a - pN3a).
- Breast cancer histopathologic grade (G1=low, G2=intermediate, G3=high, not evaluable/missing)
- Region (Japan, Other)
- Global Region (North America, Eastern Europe, Western Europe and Australia, Asia, and Other)

Other covariates reported in the literature or from other ongoing Amgen studies may be considered in the analysis as appropriate at the time of analysis.

## **5. Definitions**

### **5.1 Basic Definition**

#### **Investigational Product (IP)**

IP for this study refers to denosumab or placebo.

#### **IVRS**

An IVRS (interactive voice response system) will be used in this study to assign eligible subjects to randomized treatment, manage IP supply at the site, and unblind subjects as needed.

#### **Postmenopausal Subset**

This subset includes all randomized subjects who were postmenopausal at enrollment.

In this study, postmenopausal status is defined as follows:

- Having undergone a bilateral oophorectomy
- Age  $\geq$  60 years
- Age 45 to 59 years meeting one of the criteria, ie, either amenorrhea > 12 months with an intact uterus and  $\geq$  1 intact ovary(ies); or amenorrhea for  $\leq$  12 months and FSH and estradiol in the postmenopausal range

Note that patients who have received adjuvant or neoadjuvant chemotherapy must have met at least 1 of the above criteria for postmenopausal status prior to that chemotherapy.

### **5.2 Study Points of Reference**

#### **Study Day 1**

The date of the first IP administration or the date of enrollment for subjects who are not administrated any dose of IP.

#### **Study Day**

The number of days from the study day 1 to a date of interest, inclusive:

Study day = (date of interest – study day 1) + 1.

#### **Analysis Visit Window**

Per protocol, all tests and procedures will be performed within a time window to allow for variations in scheduling. The analysis visit windows defined in [Section 12.1.2](#) will be used to assign measurements to the appropriate nominal visit for the analyses.

### **5.3 Study Dates**

#### **Informed Consent Date**

The informed consent date is the date on which the subject signs the informed consent.

### **Enrollment (Randomization) Date**

The date on which a subject is assigned to one of the treatment arms through the IVRS system.

### **First Dose Date**

The first dose date is the date on which a subject is administered the first non-zero dose of IP, required by protocol to be  $\leq 8$  calendar days after randomization to the study.

### **Last Dose Date**

The date on which a subject is administered her final non-zero dose of IP.

### **Primary Analysis Data Cut-off Date**

The date, defined by Amgen, on which all enrolled subjects would have had the opportunity to complete 5 years of treatment from study day 1. The primary analysis data cut-off date for this study is assigned to August 31, 2017.

### **End of Treatment Phase (EOTP) Date for a Subject**

The end of treatment phase date is the date recorded on the End of Treatment Phase page of the Case Report Form (CRF) for an enrolled subject.

### **End of Study Date for a Subject**

The end of study date is the date recorded on the End of Study page of the CRF for an enrolled subject.

## **5.4 Study Time Intervals**

### **Screening period**

The screening period for a subject is defined as the time from the informed consent date to the day of enrollment. This period is required by protocol to be  $\leq 4$  weeks.

### **On-Study Period (up to the primary analysis data cut-off date)**

For the primary analyses, the on-study period for an enrolled subject is defined as the time from the day of enrollment to the end of study date on the CRF (for subjects who terminated study before the primary analysis data cut-off date) or the primary analysis data cut-off date, whichever is earlier.

## **5.5 Subject Dispositions**

### **Enrolled**

Individuals are considered enrolled if they have been assigned a randomization number and they have a non-missing date of randomization. Enrolled individuals are referred to as “subjects”.

### **On-Study (up to the primary data cut-off date)**

For the primary analyses, subjects are considered on-study during the period between the date of enrollment and the end of study date on the CRF or the primary data analysis cut-off date, whichever comes first.

### **Exposed to Investigational Product**

Subjects are considered as exposed to IP if they have received one or more non-zero doses of IP.

## **5.6 Study-specific Definitions**

### **Onsite Clinical Visit**

In the treatment phase, onsite clinical visits are determined as the visits at which physical exam and/or vital signs and/or ECOG have been assessed or IP dosing was performed. During LTFU, onsite clinical visits are determined as the visits at which the type of contact is documented as Clinical Visit or Another Physician on the Follow-up Survival Status CRF page.

### **Solitary Lesion (Via Central Imaging Analysis)**

Subjects are considered to have a solitary lesion if there is only a single lesion, bone metastasis or extraosseous disease (EOD) lesion, identified by the central imaging vendor across all body sites. If the single lesion is inside the bone, the subject is considered to have a solitary bone metastasis lesion. Otherwise, the subject is considered to have a solitary EOD lesion.

### **Bone Metastasis**

Bone metastasis must be confirmed by central imaging analysis or by biopsy, and it includes any of the following:

- Bone metastasis lesion that is confirmed via central imaging analysis and is not a solitary lesion
- Solitary bone metastasis lesion via central imaging analysis [see definition of “solitary lesion (via central imaging analysis)” above] confirmed by the investigator with a positive biopsy unless biopsy is medically contra-indicated
- Bone metastasis determined by the investigator and confirmed via positive biopsy

Any positive skeletal scintigraphy (‘bone scan’) (per central imaging analysis) must be confirmed using X-ray, computerized tomography (CT), magnetic resonance imaging (MRI), or biopsy evidence of lesions consistent with bone metastasis. Evidence of disseminated tumor cells in bone marrow is not sufficient for determination of disease

recurrence. Development of new primary malignancy in bone will not be considered as bone metastasis.

### **Extrasosseous Disease (EOD)**

EOD must be confirmed by the central imaging analysis or by biopsy/cytology, and it includes any of the following

- EOD lesion that is confirmed via central imaging analysis and is not a solitary lesion
- Solitary EOD lesion via central imaging analysis [see definition of “solitary lesion (via central imaging analysis)” above] confirmed by the investigator with a positive biopsy or cytology unless biopsy/cytology is medically contra-indicated
- EOD determined by the investigator and confirmed via positive biopsy or cytology. Cytology or biopsy is required for diagnosis of malignant effusion (eg, pleural effusion, pericardial effusion, or ascites) or meningeal carcinomatosis (cerebrospinal fluid).

Local-regional findings (see definition below) prior to surgical ( $\pm$  radiation) treatment with curative intent are not considered as EOD. For a subject with local-regional findings after the surgical ( $\pm$  radiation) treatment with curative intent, these findings will be considered as EOD (local-regional disease recurrence) if at least one of them is confirmed by positive biopsy or cytology. For a neoadjuvant subject without surgical ( $\pm$  radiation) treatment with curative intent, all her local-regional findings will be considered as EOD (local-regional disease recurrence) if at least one of these findings is confirmed by positive biopsy or cytology.

Development of non-breast cancer new primary malignancy will not be considered as EOD.

### **Local-Regional Findings**

Defined as potential development of breast cancer in any soft tissue or skin of the ipsilateral chest wall; or in the regional lymph nodes, including ipsilateral axillary, ipsilateral internal mammary, ipsilateral supraclavicular, and/or ipsilateral infraclavicular lymph nodes, and/or in the soft tissue of the ipsilateral axilla. Specifically, the local body sites include breast and non-bone chest wall. The regional body sites include axillary node, axillary non-nodal, infraclavicular nodes, internal mammary node, intramammary node, supraclavicular nodes, and subpectoral node.

### **Distant Disease Recurrence**

Distant disease recurrence is defined as the development of breast cancer in any distant site (including the skin, subcutaneous tissue [other than skin or subcutaneous tissue

areas identified above], and the contralateral breast, within others) and/or distant lymph nodes (ie, any lymph nodes other than local or regional lymph nodes as defined above). In a word, distant disease recurrence includes all bone metastasis as well as any EOD other than local-regional disease recurrence defined above. Development of non-breast cancer new primary malignancy will not be considered as distant disease recurrence.

### **Breast Cancer Disease Recurrence**

Breast cancer disease recurrence includes bone metastasis and EOD defined above.

### **Bone Metastasis [including investigator reported events]**

Bone metastasis [including investigator reported events] includes bone metastasis defined above and clinically diagnosed breast cancer recurrence in bone, specifically includes any of the following

- Bone metastasis lesion that is confirmed via central imaging analysis and is not a solitary lesion
- Solitary bone metastasis lesion via central imaging analysis [see definition of “solitary lesion (via central imaging analysis)” above] confirmed by the investigator with a positive biopsy unless biopsy is medically contra-indicated
- Bone metastasis determined by the investigator via any diagnostic procedure

Any positive skeletal scintigraphy (“bone scan”) (per central imaging analysis) must be confirmed using X-ray, CT, or MRI. Evidence of disseminated tumor cells in bone marrow is not sufficient for determination of disease recurrence. Development of new primary malignancy in bone will not be considered as bone metastasis.

### **Extraosseous Disease (EOD) [including investigator reported events]**

EOD [including investigator reported events] includes EOD defined above and clinically diagnosed breast cancer recurrence at a body location outside the bone, specifically includes any of the following

- EOD lesion that is confirmed via central imaging analysis and is not a solitary lesion
- Solitary EOD lesion via central imaging analysis [see definition of “solitary lesion (via central imaging analysis)” above] confirmed by the investigator with a positive biopsy or cytology unless biopsy/cytology is medically contra-indicated
- EOD determined by the investigator via any diagnostic procedure

Local-regional findings (see definition above) prior to surgical ( $\pm$  radiation) treatment with curative intent are not considered as EOD [including investigator reported events]. For a subject with local-regional findings after the surgical ( $\pm$  radiation) treatment with curative intent, these findings will be considered as EOD (local-regional disease recurrence)

[including investigator reported events] if at least one of them is determined by the investigator via any diagnostic procedure. For a neoadjuvant subject without surgical ( $\pm$  radiation) treatment with curative intent, all her local-regional findings will be considered as EOD (local-regional disease recurrence) [including investigator reported events] if at least one of these findings is determined by the investigator via any diagnostic procedure.

Development of non-breast cancer new primary malignancy will not be considered as EOD [including investigator reported events]

### **Breast Cancer Disease Recurrence [including investigator reported events]**

Breast cancer disease recurrence [including investigator reported events] includes bone metastasis [including investigator reported events] and EOD [including investigator reported events] defined above.

### **Fracture (Vertebral or Non-vertebral Fracture, Prior to Bone Metastasis)**

Only fractures occurring before 12 weeks prior to the first bone metastasis are considered. Fractures (vertebral or non-vertebral) includes fractures, which are confirmed by central imaging analysis, occurring at any site except skull, facial, mandible, metacarpus, finger phalanges, and toe phalanges. In addition, fractures associated with high trauma severity (ie, fall from higher than about 20 inches, or severe trauma other than a fall) will be excluded.

### **Skeletal Related Event (SRE)**

A SRE is defined as one of the following: pathologic fracture, radiation therapy to bone, surgery to bone, or spinal cord compression. All such events that occur from 12 weeks prior to the first confirmed bone metastasis are considered SRE in this study.

### **Pathologic Fracture**

Pathologic fractures are those new bone fractures that occur spontaneously in subjects with bone metastases and are not a result of severe trauma (ie, fall from higher than about 20 inches, or severe trauma other than a fall). Pathologic fractures are identified or confirmed by the central imaging vendor. The severity of the trauma will be determined by the investigator on the CRF. If a fracture is reported by central imaging vendor but not in the CRF it will be considered as not due to severe trauma.

### **Radiation Therapy to Bone**

Radiation therapy to bone in subjects with bone metastases includes radiation for pain control (including use of radioisotopes), to treat or prevent pathologic fractures, or to

treat or prevent spinal cord compression. These events will be captured on the CRF. Radiation may be administered over a period of time. Only the start date of radiation will be relevant when counting skeletal events.

### **Surgery to Bone**

Surgery to bone in subjects with bone metastases includes procedures to set or stabilize a pathologic fracture, or to prevent an imminent pathologic fracture or spinal cord compression. These events will be captured on the CRF.

### **Spinal Cord Compression**

Spinal cord compression events in subjects with bone metastases must be confirmed by the central imaging vendor using appropriate radiographic imaging (eg, myelogram, MRI or CT scans), and not due to trauma per investigator's evaluation.

### **Hypercalcemia of Malignancy**

Clinical events of hypercalcemia of malignancy will be collected on the CRF throughout the study and during LTFU prior to primary analysis data cut-off date. Hypercalcemia of malignancy is defined in this study as a serum calcium value (albumin-adjusted if necessary) of CTCAE version 3.0 grade 2 or greater (ie, > 11.5 mg/dL) during the treatment phase, and CTCAE grade 2 or greater events recorded on the Calcium Abnormalities CRF page during LTFU.

### **Breast Density**

Breast density will be determined based on standard mammography imaging, and assessed according to the Breast Imaging-Reporting and Data System (BI-RADS) classification of overall breast composition as follows:

- The breast is almost entirely fat (< 25% glandular)
- There are scattered fibroglandular densities (approximately 25 – 50% glandular)
- The breast tissue is heterogeneously dense, which could obscure detection of small masses (approximately 51 – 75% glandular)
- The breast tissue is extremely dense. This may lower the sensitivity of mammography (> 75% glandular).

### **Pathological Complete Response**

There has not been a uniform definition of pathological complete response. In this study, pathological complete response following the completion of surgery for breast cancer for neo-adjuvant subjects is assessed per investigators' justification based on the evaluation performed by the local pathologist. A complete pathologic response is considered if no invasive or noninvasive residual disease was found in the breast or axillary lymph nodes.



### **Residual Invasive Tumor Size (RITS)**

There has not been a uniform definition of RITS. In this study, RITS following the completion of surgery for breast cancer for neo-adjuvant subjects is assessed per investigators' justification based on the evaluation performed by the local pathologist. A RITS "0 mm" is considered if no residual invasive tumor was recorded in the breast or axillary lymph nodes.

## **5.7 Derived Variables**

### **Baseline Value**

Baseline value is the latest recorded measurement on or prior to the day of the first dose of IP. If a subject doesn't receive IP, baseline is the latest recorded measurement on or prior to the enrollment date.

### **Change from Baseline Value**

The arithmetic difference between a value of interest and a baseline value: Change from baseline value = (value of interest - baseline value)

### **Percent Change from Baseline Value**

The ratio of the arithmetic difference between a value of interest and the baseline value to the baseline value multiplied by 100: Percent change from baseline value = [(value of interest - baseline value) / baseline value] \* 100.

### **Subject Incidence Rate**

The subject incidence rate for a given event is defined as the number of subjects with one or more reported occurrence of the event divided by the number of subjects who have the opportunity to report the event.

### **Subject Age at Study Entry**

Age at study entry is defined as the number of whole years from a subject's birth date to the enrollment date as calculated in the EDC RAVE database.

### **Time to Event for Efficacy**

Time interval (days) from the randomization date to the date of occurrence of the event or censorship during the given period:

Time interval = (date of occurrence of the event or censorship – randomization date) +1

### **Time to Event for Safety**

Time interval (days) from the first dose date to the date of occurrence of the event or censorship during the given period:

Time interval = (date of occurrence of the event or censorship – first dose date) +1

### **Last Assessment Date (Via Central Imaging Analysis) for Bone Metastasis**

Defined as the exam date of the latest imaging record among all the following imaging records from the central imaging analysis. If no such record exists on study, the last assessment date (via central imaging analysis) for bone metastasis is defined as the randomization date.

- Whole-body planar radioisotope bone scan
- Whole-body single-photon emission CT (SPECT) scan

### **Last Assessment Date (Via Central Imaging Analysis) for Disease Recurrence**

Defined as the exam date of the latest imaging record among all the following imaging records from the central imaging analysis. If no such record exists on study, the last assessment date (via central imaging analysis) for disease recurrence is defined as the randomization date.

- Whole-body planar radioisotope bone scan
- Whole-body single-photon emission CT (SPECT) scan
- CT of the chest
- CT of the abdomen
- MRI of the chest
- MRI of the abdomen

### **Bone Metastasis-free Survival (BMFS) Time**

Time interval (days) from the randomization date to the date of first occurrence of bone metastasis (ie, first documented evidence per central imaging analysis or per the investigator via confirmation procedure of biopsy) or death from any cause, whichever comes first. Subjects last known to be alive, who have not experienced bone metastasis, are censored at their last assessment date (via central imaging analysis) for bone metastasis defined above, or at the primary analysis data cut-off date, whichever comes first. Subjects who had first occurrence of bone metastasis before randomization will be censored at their randomization date.

### **Disease-free Survival (DFS) Time**

Time interval (days) from the randomization date to the date of first observation of disease recurrence (ie, first documented evidence per central imaging analysis or per the investigator via confirmation procedure of biopsy or cytology) or death from any cause, whichever comes first. Subjects last known to be alive, who have not experienced recurrence of disease, are censored at their last assessment date (via central imaging analysis) for disease recurrence defined above, or at the primary analysis data cut-off date, whichever comes first. Subjects who had first disease recurrence before randomization will be censored at their randomization date.

### **Overall Survival (OS) Time**

Time interval (days) from the randomization date to the date of death from any cause. Subjects last known to be alive are censored at their last contact date, or at the primary analysis data cut-off date, whichever comes first.

### **Distant Recurrence-free Survival (DRFS) Time**

Time interval (days) from the randomization date to the date of first observation of distant disease recurrence or death from any cause, whichever comes first. Subjects last known to be alive, who have not experienced distant disease recurrence, are censored at their last assessment date (via central imaging analysis) for disease recurrence defined above, or at the primary analysis data cut-off date, whichever comes first. Subjects who had first distant recurrence before randomization will be censored at their randomization date.

### **Time to First Bone Metastasis (excluding death)**

Time interval (days) from the randomization date to the date of first observation of bone metastasis (ie, first documented evidence per central imaging analysis or per the investigator via confirmation procedure of biopsy). Subjects, who have not experienced bone metastasis, are censored at their last assessment date (via central imaging analysis) for bone metastasis defined above, or at the primary analysis data cut-off date, whichever comes first. Subjects who had first occurrence of bone metastasis before randomization will be censored at their randomization date.

### **Time to Bone Metastasis as Site of First Recurrence (excluding death)**

Time interval (days) from the randomization date to the date of first observation of bone metastasis as site of first disease recurrence. Subjects, who have not experienced bone metastasis as site of first recurrence, are censored at the time of first EOD recurrence (for cause-specific analysis), or at their last assessment date (via central imaging

analysis) for bone metastasis defined above, or at the primary analysis data cut-off date, whichever comes first. Subjects who had first disease recurrence before randomization will be censored at their randomization date.

#### **Time to Disease Recurrence**

Time interval (days) from the randomization date to the date of first observation of disease recurrence, excluding death. Subjects, who have not experienced recurrence of disease, are censored at their last assessment date (via central imaging analysis) for disease recurrence defined above, or at the primary analysis data cut-off date, whichever comes first. Subjects who had first disease recurrence before randomization will be censored at their randomization date.

#### **Time to Distant Recurrence**

Time interval (days) from the randomization date to the date of first observation of distant recurrence, excluding death. Subjects, who have not experienced distant recurrence, are censored at their last assessment date (via central imaging analysis) for disease recurrence defined above, or at the primary analysis data cut-off date, whichever comes first. Subjects who had first distant recurrence before randomization will be censored at their randomization date.

#### **Time to First On-study Fracture (Prior to Bone Metastasis)**

Time interval (days) from the randomization date to the date of first fracture (vertebral or non-vertebral fracture) prior to the development of bone metastasis (up to 12 weeks prior to the date of first diagnosis of bone metastasis). Subjects, who have not experienced a fracture on study, are censored at 12 weeks before the time of development of bone metastasis, at their last contact date, or at the primary analysis data cut-off date, whichever comes first.

#### **Time to First SRE**

Time interval (days) from the randomization date to the date of first on-study SRE (pathologic fracture, radiation therapy to bone, surgery to bone, or spinal cord compression) following the development of bone metastasis (including 12 weeks prior to the date of first diagnosis of bone metastasis). Subjects who have not experienced an SRE on study are censored at their last contact date, or at the primary analysis data cut-off date, whichever comes first.

### **Time to First SRE or Hypercalcemia**

Time interval (days) from the randomization date to the date of first on-study SRE (pathologic fracture, radiation therapy to bone, surgery to bone, or spinal cord compression) or hypercalcemia following the development of bone metastasis (including 12 weeks prior to the date of first diagnosis of bone metastasis). Subjects, who have not experienced a SRE or hypercalcemia on study, are censored at their last contact date, or at the primary analysis data cut-off date, whichever comes first.

### **Time to First On-study Symptomatic Bone Metastasis**

Time interval (days) from the randomization date to the date of first on-study observation of bone metastasis which is accompanied by symptom at the time of detection. Subjects who have not experienced symptomatic bone metastasis are censored at their first on-study asymptomatic bone metastasis date (for cause-specific analysis), at their last assessment date (via central imaging analysis) for bone metastasis defined above, or at the primary analysis data cut-off date, whichever comes first. Subjects who had first symptomatic bone metastasis before randomization will be censored at their randomization date.

### **Time to $\geq$ 2-point Increase from Baseline in “Worst” Pain Score**

Time interval (days) from the randomization date to the date of first observation of  $\geq$  2-point increase from baseline in “Worst” pain score. Subjects, who have not experienced the event of interest, are censored at the later of the randomization date and the last BPI-SF assessment date, or at the primary analysis data cut-off date, whichever comes first.

### **Time to $\geq$ 2-point Increase from Baseline in “Worst” Pain Score (After 6 Months Post Randomization)**

Time interval (days) from the randomization date to the date of first observation of  $\geq$  2-point increase from baseline in “Worst” pain score occurred after 6 months post randomization. Subjects, who have not experienced the event of interest after 6 months post randomization, are censored at the later of 6 months post randomization and the last BPI-SF assessment, or at the primary analysis data cut-off date, whichever comes first.

### **Time to > 4-point in “Worst” Pain Score**

Time interval (days) from the randomization date to the date of first observation of > 4-point in “Worst” pain score. Subjects, who have not experienced the event of

interest, are censored at the later of the randomization date and the last BPI-SF assessment, or at the primary analysis data cut-off date, whichever comes first.

### **Time to > 4-point in “Worst” Pain Score (After 6 Months Post Randomization)**

Time interval (days) from the randomization date to the date of first observation of > 4-point in “Worst” pain score occurred after 6 months post randomization. Subjects, who have not experienced the event of interest after 6 months post randomization, are censored at the later of 6 months post randomization and the last BPI-SF assessment, or at the primary analysis data cut-off date, whichever comes first.

### **Analgesic Score**

Analgesic score is defined as follows.

- 0 = No analgesics
- 1 = Non-opioid analgesics
- 2 = Weak opioids (eg, meperidine, codeine, tramadol)
- 3 = Strong opioids  $\leq$  75 mg oral morphine equivalent (OME) per day
- 4 = Strong opioids with > 75 - 150 mg OME per day
- 5 = Strong opioids with > 150 - 300 mg OME per day
- 6 = Strong opioids with > 300 - 600 mg OME per day
- 7 = Strong opioids with > 600 mg OME per day

Analgesic score will be used as numerical value. For guidance on analgesic scoring derivation process and analgesic assumptions, please refer to User Manual, Opioid Analgesic Quantification Algorithm for Denosumab Oncology Skeletal Related Event Clinical Studies. The User Manual will be updated before unblinding of the study for primary analysis, and kept on file at Amgen.

### **Subject Years on Study for Safety Analysis**

The sum of days on study ([30 days after the last dose of IP for subjects who discontinued/completed IP or the primary analysis data cut-off date, whichever comes first] – first dose date +1) divided by 365.25.

### **Subject year-adjusted Adverse Event Rate**

The number of adverse events reported in the given time period divided by total subject years on study during the period.

**The visit when  $\geq 30\%$  of subjects dropped out due to death, disease progression or consent withdrawn**

This is used in some PRO analyses and is defined as the first visit when  $\geq 30\%$  of subjects dropped out due to death, disease progression or consent withdrawn among those who have had an opportunity to reach that visit.

**6. Analysis Data Sets**

**6.1 Data Subsets**

**6.1.1 Full Analysis Set (FAS)**

The FAS includes all subjects who are randomized. Subjects in the FAS will be analyzed according to their original treatment assignment, regardless of treatment received. Data collected on subjects in this subset during the period from randomization to the primary analysis data cut-off date will be included, regardless of compliance with the protocol.

**6.1.2 Per Protocol Set**

The per protocol set is defined as all subjects with a protocol-defined diagnosis, no major protocol violations, no disease recurrence before study day 1, and who received at least 1 dose of investigational product. Subjects with deviations from the following protocol-specified eligibility criteria will be excluded from the per-protocol set:

1. Inclusion criteria:

- Histologically confirmed, AJCC stage II or III breast cancer
- High risk of breast cancer recurrence, defined as documented evidence of one or more of the following criteria:
  - Biopsy evidence of breast cancer in regional LN (node positive disease)
  - *Nodal micrometastases only are not considered node positive*
  - Tumor size  $> 5$  cm (T3) or locally advanced disease (T4)
- Documented pathological evaluation of the breast cancer for hormone receptor (estrogen receptor [ER] and progesterone receptor [PR]) status and HER-2 status
- For subjects receiving adjuvant therapy only:
  - If the subject did not have complete resection of the primary tumor by the time of randomization, definitive treatment must be planned to be completed within approximately 12 months of randomization
  - Time between definitive surgery and randomization must be  $\leq 12$  weeks
  - If the subject had node-positive disease but did not undergo treatment of axillary LN with curative intent prior to randomization, definitive treatment must be planned to be completed within approximately 12 months of randomization
  - Subjects must not have received prior neoadjuvant treatment
  - *Endocrine treatment for less than 30 days prior to surgery is not considered prior neoadjuvant treatment*

- For subjects receiving neoadjuvant therapy only:
  - Time between start of neoadjuvant treatment and randomization must be  $\leq 8$  weeks
  - Subjects must be scheduled to undergo definitive treatment (including surgery and/or radiotherapy) with curative intent within approximately 12 months of starting neoadjuvant treatment
- Serum calcium or albumin-adjusted serum calcium not less than 1.75 mmol/L (7.0 mg/dL)
- Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1
- Written informed consent before any study-specific procedure is performed

2. Exclusion criteria:

- Prior or current evidence of any metastatic involvement of any distant site
- History of breast cancer (other than ductal carcinoma in situ [DCIS] or lobular carcinoma in situ [LCIS]) prior to the current diagnosis
- Osteoporosis requiring treatment at the time of randomization or treatment considered likely to become necessary within the subsequent six months
- Any prior or synchronous malignancy (other than breast cancer), except:
  - Malignancy treated with curative intent and with no evidence of disease for  $\geq 5$  years prior to enrollment and considered to be at low risk for recurrence by the treating physician
  - Adequately treated non-melanoma skin cancer or lentigo maligna without evidence of disease
- Active infection with Hepatitis B virus or Hepatitis C virus
- Known infection with human immunodeficiency virus (HIV)
- Prior history or current evidence of osteomyelitis/ osteonecrosis of the jaw
- Active dental or jaw condition which requires oral surgery
- Non-healed dental or oral surgery
- Prior or current IV bisphosphonate administration
- Prior administration of denosumab
- Subject currently is enrolled in or has not yet completed at least 30 days since ending other investigational device or investigational drug study(s), or subject is receiving other investigational agent(s)
- Subject has known sensitivity to any of the products to be administered during the study (eg, mammalian derived products, calcium, or vitamin D)

Time to event endpoints will be defined from the 1<sup>st</sup> dose date to the event of interest or censoring time. For a subject who has not experienced the event of interest, she is censored at the last assessment date (per definition of the time to event endpoint in [Section 5.7](#)), at the earliest violation date (if she has experienced any of the following violations), or at the primary analysis data cut-off date, whichever comes first. For a



subject who has experienced the event of interest after any of the following violations, they will be censored at the earliest violation date, or at the primary analysis data cut-off date, whichever comes first.

- Receiving a dose(s) of investigational product not matching the subject's randomized treatment group. The violation date is the first date that the incorrect investigation product was administered.
- Receiving one treatment regimen of proscribed medication or therapy on study. Proscribed therapy includes administration of IV bisphosphonates, denosumab available through commercial sources (eg, XGEVA™, Prolia®) or oral bisphosphonates for the adjuvant/neoadjuvant treatment of breast cancer, or any unapproved investigational product(s) or investigational device(s) while subject are on study treatment. The violation date is the first date that proscribed medication or therapy was received.
- Not receiving investigational product for 3 consecutive doses due to reasons other than investigational product related CTCAE grade 3 or 4 adverse event, or suspected atypical femoral fracture (AFF) or osteonecrosis of the jaw (ONJ), or planned invasive dental procedure /healing from invasive dental procedure. The violation date is 98 days after the time of last previously received dose.

Subjects in the per-protocol subset will be analyzed according to the randomized treatment assignment and the actual stratum.

### **6.1.3 Safety Analysis Set**

This subset includes all subjects who are randomized and receive  $\geq 1$  dose of investigational product. Subjects in this subset will be analyzed according to their actual treatment received; subjects who received  $\geq 1$  dose of denosumab will be analyzed in the denosumab treatment group regardless of the randomized treatment assigned.

### **6.1.4 Pharmacokinetic (PK) Analysis Set**

The PK analysis set is the subset for analyses of denosumab serum concentration levels. This subset includes subjects who enrolled into the pharmacokinetics portion of the trial and who also received at least one dose of denosumab.

## **6.2 Subset for Interim Analysis**

There is 1 formal interim analysis after approximately 367 subjects have developed bone metastasis or died, and approximately 555 subjects have develop any disease recurrence or died, whichever occurs last. This interim analysis may be conducted as soon as approximately 6 months after the sufficient event number has been reached for either BMFS or DFS in the full study population, even if the event number for the other endpoint has not been reached. All subjects enrolled by the time of the interim analysis data cut-off dates will be included in the interim analysis.

### 6.3 Subgroup Analyses

The primary and key secondary efficacy endpoints will be analyzed in the following subgroups:

- Breast cancer therapy / Lymph node status (neo-adjuvant therapy / any LN status, adjuvant therapy / LN negative, and adjuvant therapy / LN positive)
- Breast cancer therapy (neo-adjuvant, adjuvant)
- Hormone receptor (ER/PR) status (positive, negative)
- HER-2 status (positive, negative)
- Molecular subtypes of breast cancer (hormone receptor positive and HER-2 positive, hormone receptor positive and HER-2 negative, hormone receptor negative and HER-2 positive, hormone receptor negative and HER-2 negative)
- Age (<50 years , ≥ 50 years)
- Menopausal status (premenopausal, postmenopausal)
- Postmenopausal status (>5 years since menopause, other [including premenopause, ≤5 years since menopause, and unknown])
- Region (Japan, Other)
- Global Region (North America, Eastern Europe, Western Europe and Australia, Asia, and Other)
- Race (Caucasian, or Non-Caucasian)

These subgroups, except for stratification variables, will be re-examined and appropriately re-categorized before unblinding. The treatment by subgroup interaction will be examined and tested as described in [Section 9.5.1.5](#).

The safety endpoints including all AEs and SAEs, will be analyzed in the following subgroup:

- Menopausal status (premenopausal, postmenopausal)

### 7. Interim Analysis and Early Stopping Guidelines

A DMC external to Amgen and external to Daiichi Sankyo Co., Ltd. will be formed with members consisting of individuals chosen for their expertise in oncology, bone disease, and in statistical methods in clinical trials. The primary role of the external DMC is to protect the safety of the study participants throughout the study duration. They will accomplish this by periodically reviewing safety and efficacy analysis results in order to assess the risk-benefit profile for subjects enrolled in this study. Access to both safety and efficacy data will provide the DMC a broad perspective with which to generate recommendations for appropriate actions. The external individuals serving on the DMC will have access to subjects' individual treatment assignments. To minimize the potential

introduction of bias, these individuals will not have direct contact with the study site personnel or subjects. Selected Amgen staff may serve as liaisons to the external DMC, but will not be voting members, and will not be unblinded to the results. Members of this external DMC will include, at a minimum, physicians external to Amgen and Daiichi Sankyo Co., Ltd., and appropriate statistical representation external to Amgen and Daiichi Sankyo Co., Ltd.

The data for review are outlined in the DMC charter and agreed to in advance by the DMC members. Unblinded reports for review will be generated by an independent CRO that is external to Amgen as well. The external DMC will convene approximately twice yearly, and the start date will depend on subject accrual rates. It is recognized that the DMC may feel ethically compelled to recommend early stopping in the event of overwhelming efficacy. The stopping rules are defined as follows:

1. There is 1 formal interim analysis after approximately 367 subjects have developed bone metastasis or died, and approximately 555 subjects have developed any disease recurrence or died, whichever occurs last. This analysis may be performed as soon as approximately 6 months after the sufficient event number has been reached for either BMFS or DFS in the full study population, even if the event number for the other endpoint has not been reached. The critical p-values for rejecting the null or alternative hypothesis are listed in [Table 11](#). The multiplicity from testing DFS in the full study population and in the postmenopausal subset simultaneously will be adjusted based on a Hochberg procedure ([Sakamaki, 2013](#); [Ye et al., 2012](#)). The study could potentially be stopped if both BMFS and DFS in the full study population cross their corresponding efficacy boundaries, or if BMFS crosses the futility boundary at the interim analysis. The other secondary endpoints OS and DRFS will only be tested when the study is stopped or at the primary analysis; the same procedure as for BMFS will be used to adjust for multiplicity between the interim and primary analysis for each of these endpoints.

**Table 11. Decision Rule (1-sided) at Interim/Primary Analyses**

Analysis	Reject H0	Reject H1
Interim Analysis 1*	<u>BMFS</u> : $P < 0.0015$ <u>DFS in the full study population</u> : <ul style="list-style-type: none"> <li>• <math>P &lt; 0.0004</math> or</li> <li>• <math>P &lt; 0.0015</math> for both the full study population and the subset</li> </ul> <u>DFS in the subset</u> : <ul style="list-style-type: none"> <li>• <math>P &lt; 0.0004</math> or</li> <li>• <math>P &lt; 0.0015</math> for both the full study population and the subset</li> </ul>	$P > 0.9333$ for BMFS
Primary Analysis**	<u>BMFS</u> : $P < 0.0247$ <u>DFS in the full study population</u> : <ul style="list-style-type: none"> <li>• <math>P &lt; 0.0125</math>; or</li> <li>• <math>P &lt; 0.0249</math> for the full study population and <math>P &lt; 0.0248</math> for the subset</li> </ul> <u>DFS in the subset</u> : <ul style="list-style-type: none"> <li>• <math>P &lt; 0.0125</math>; or</li> <li>• <math>P &lt; 0.0249</math> for the full study population and <math>P &lt; 0.0248</math> for the subset</li> </ul>	

\*At the interim analysis 1, the study could potentially be stopped if both BMFS and DFS in the full study population cross their corresponding efficacy boundaries, or if BMFS crosses the futility boundary as shown in the table

\*\*Boundaries are based on the estimated number of events and will be recalculated based on the actual observed number of events.

2. At other interim evaluations, the DMC will focus on safety, and will review efficacy data only to balance the risk-benefit assessment. At these analyses, the guideline is that p-values for both BMFS and DFS of less than 0.0005 will be used as evidence of overwhelming efficacy to stop the study.
3. At the primary analysis that is time-driven, the exact number of events for the primary and secondary endpoints are not fixed and the final alpha boundaries can potentially vary. Table 11 includes the boundaries calculated according to the estimated total numbers of events. In order to strictly control the overall type I error rate, the final boundaries will be updated based on the actual number of events achieved and the alpha levels that have been spent at the interim analysis (Proschan et al, 2006).

As described in the DMC charter, recommendations of the DMC will be communicated by the DMC Chair to senior management at Amgen for discussion on how to proceed. Amgen senior management will be responsible for making decisions for further dissemination of the recommendations at that time.

## **8. Data Screening and Acceptance**

### **8.1 General Principles**

The objective of data screening is to assess the quantity, quality, and statistical characteristics of the data relative to the requirements of the planned analyses. Any questionable values or situations will be reported to Clinical Data Management (CDM) for review or confirmation.

### **8.2 Data Handling and Electronic Transfer of Data**

Amgen's Data Management department will provide all data to be used in the planned analyses. This study will use the Electronic Data Capture (EDC) database, RAVE. All laboratory values will be electronically transferred from the central laboratory to the Amgen database. The central imaging vendor will assess the images submitted by investigators (eg, Mammography, Bone Scan, CT/MRI) and electronically transfer the data to Amgen.

### **8.3 Handling of Missing and Incomplete Data**

Subjects may have missing data points for a variety of reasons. In general, data may be missing due to subject's early withdrawal from study, a missed visit, or non-evaluability of an endpoint at a particular point in time. For the study of time-to-event endpoints, subjects completing the study and not experiencing an event of interest are considered censored to the event (see [Section 5.7](#)). The procedures outlined below describing what will be done when data are missing may be refined during the blind review of the data.

#### **8.3.1 Missing Dates**

Incomplete event start dates and concomitant medications start or stop dates will be imputed as described in [Table 12](#). If the start date is complete missing, for AE, default to study day 1 if the indicator of "Did event start before the first dose of investigational product" is checked as "No"; for concomitant medication, assume it started before enrollment. If the stop date and the flag for continuing are both missing for an on-study event or medication, assume the event or medication stopped after the end of study date. Partial dates will be listed as is on the listings.

**Table 12. Imputation Rules for Incomplete Dates**

	Missing	Imputation	Exception
Start date (AE, and concomitant medication)	Day	01	Default to Study Day 1 if an event starts the same year and month as Study Day 1
	Day / Month	01JAN	Default to Study Day 1 if an event started the same year as Study Day 1
Start date (disease recurrence, and SRE)	Day	01	Default to randomization date if an event starts the same year and month as randomization date
	Day / Month	01JAN	Default to randomization date if an event started the same year as randomization date
Stop date (concomitant medication only)	Day	Last day of the month	Default to the End of Study Date or the primary analysis data cut-off date for ongoing subjects if the concomitant medication stopped the same year and month as the End of Study Date or primary analysis data cut-off date for ongoing subjects respectively
	Day / Month	31DEC	Default to the End of Study Date or the primary analysis data cut-off date for ongoing subjects if the concomitant medication stopped the same year as the End of Study Date or primary analysis data cut-off date for ongoing subjects respectively

For historical events (eg, date of initial diagnosis of breast cancer, date of last menstrual period) that will be used to derive baseline covariates, the imputation rules are as follows: if the day is missing, default to day 15; if both month and day are missing, default to July 1. If the imputed date is on or after the randomization date, default to randomization date minus 1. Missing years will not be imputed under any conditions.

If a death date is incomplete and missing only the day field, the partial death date (also end of treatment phase or end of study date if death is the end-of-treatment-phase or end-of-study reason) will be imputed as follows: if the latest assessment for this subject is before the month of the death, it will be imputed as the first day of the month; if the latest assessment is within the same month as the death, it will be imputed using this latest assessment date.

For PRO analysis, if the date on the PRO administration page is missing, the date of IP administration at the corresponding visit with the same FOLDER will be used.

### **8.3.2 Assayed Values Below or Above Quantifiable Limits of Bone Turnover Marker (BTM)**

BTM measurements that are below quantifiable limit (BQL) or above quantifiable limit (AQL) will be considered equal to the limits of quantification for numerical analyses unless explicitly noted otherwise.

### **8.3.3 Missing / Unknown Values of Covariates**

For continuous covariates, the missing / unknown values will be imputed using the mean of the pooled observed data. For categorical covariate, the missing / unknown values will be combined with the category with the most subjects if the missing / unknown rate is  $\leq 2\%$  of the pooled observed data or they will be classified into a separate category if the missing/unknown rate is  $> 2\%$  of the pooled observed data. If the covariates are also used as subgroup variables, the imputed values of subgroup variables will not be used for subgroup categorization.

### **8.3.4 Missing PRO data**

Multiple imputation is used as the imputation method for missing PRO data in this study. The imputation model includes the corresponding PRO assessments up to the visit when  $\geq 30\%$  subjects have ended treatment phase due to death, disease progression or consent withdrawn, and the covariates used in the Cox proportional hazards model described in [Section 9.5.2.3](#).

### **8.3.5 Missing Pathological Complete Response**

For neo-adjuvant subjects without receiving primary breast cancer surgery, the missing pathological completed response will be imputed as “No”.

## **8.4 Assessment of Outliers**

Descriptive statistics will be used to identify outliers in any key variables defined in [Section 4](#). For continuous variables, the univariate analysis will be conducted and box plots will be generated to identify outliers. For discrete variables, frequency summaries will be examined to identify questionable values.

Outliers due to data entry error will be corrected in the database before the before analysis snapshot and database lock. Outliers that are not due to data entry error will be included in the primary analysis, unless there is sufficient clinical justification obtained prior to unblinding to exclude them. The validity of any questionable values or outliers will be confirmed. No valid measurement may be excluded from descriptive or inferential analyses. However, sensitivity analyses may be conducted to evaluate the impact of extreme values on the data.

## 8.5 Distributional Characteristics

The assumption of proportional hazards for proportional hazards models ([Lawless, 1982](#)) will be examined by visualizing plots of log (-log (survival function)) versus time and plots of Schoenfeld residuals versus time ([Schoenfeld, 1982](#)). When there is modest deviation from the assumption, inferences in proportional hazards models are generally fairly robust.

The normality assumptions in the parametric models analyzed for continuous data will be checked. If the assumptions for the distributional characteristics of endpoints are not met, these will be described and further analyses may be carried out using transformations of endpoint values.

## 8.6 Validation of Statistical Analysis

Programs will be developed and maintained, and output will be verified in accordance with current risk-based quality control procedures.

Tables, figures, and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analysis consists of Amgen-supported versions of statistical analysis software, for example, the SAS® System 9.4.

## 9. Statistical Methods of Analysis

### 9.1 General Principles

The analyses of efficacy and safety endpoints will be based on the analysis sets defined in [Section 6.1](#). Sensitivity analyses for primary and key secondary efficacy endpoints will be performed as described in [Section 9.5.1.4](#) to test the robustness of the treatment effects observed. Two-sided p-values will be presented for all the statistical testing.

Continuous variables will be summarized using descriptive statistics, which includes n (number of non-missing observations), the mean, median, 25<sup>th</sup> and 75<sup>th</sup> percentiles, standard deviation, minimum, and maximum. The minimum and maximum will be reported using the same precision as the original measurement. The mean, median, other selected percentiles, and standard deviation will be reported to one decimal place more than the precision of the original measurement. For categorical variables, descriptive statistics include frequency and percentage.

For time to event variables, the Kaplan-Meier method will be used to estimate the quartiles (25<sup>th</sup> percentile and median) of time to event variable for each treatment group, along with 95% confidence intervals ([Kaplan and Meier, 1958](#)). Kaplan-Meier estimates



will be presented graphically for each treatment group approximately up to the time point where there are less than 50 subjects in the risk set for both treatment groups combined. Kaplan-Meier event rates with 2-sided 95% confidence intervals at year 1, year 2, year 3, year 4 and year 5 will also be estimated. The hazard ratio and its 2-sided 95% CI will be estimated using Cox proportional hazards model (Cox, 1972) with treatment group as the independent variable and stratified by the randomization stratification factors.

For analyses where stratification needs to be adjusted for, the following general principles will be followed:

- Primary analyses, which are intended to evaluate treatment effects and are using stratified models, will be based on the randomized stratum, regardless of the subject's actual value. If > 5% of all randomized subjects were stratified incorrectly for 1 or more factors, then a sensitivity analysis using the actual value should also be performed for the primary and key secondary endpoints.
- Covariate analyses, where covariates are stratification factors, will be based on subject's actual value, regardless of the stratum that they were randomized to.
- For subgroup analyses,
  - Analyses will be based on the randomized stratum.
  - If the subgroup variable is a stratification factor, sensitivity analyses according to actual value may be needed.

## 9.2 Subject Accountability

For all randomized subjects, the number and percentage achieving the planned assessments listed below will be tabulated by treatment arm:

- Subject enrollment by region, country and site
- Subjects who enrolled, who discontinued study, and their reasons for study discontinuation
- Subjects who never received IP, who received IP, who discontinued IP, and their reasons for IP discontinuation
- Subjects in each key analysis set
- Screening disposition and reasons for screen failure
- Important protocol deviations

In addition, listings will be provided for randomization assignment, important protocol deviations, protocol inclusion/exclusion deviations, manufacturing lot numbers and manufacturing lot numbers used for all enrolled subjects.

## 9.3 Demographic and Baseline Characteristics

Descriptive statistics will be tabulated for demographics, baseline characteristics and disease history as listed below by randomized treatment group using FAS.

### 9.3.1 Demographics

- Gender
- Age at study entry (year)
- Age group (18 - 64 years, 65 - 74 years, 75 - 84 years,  $\geq$  85 years)
- Age group 2 (<50 versus  $\geq$  50)
- Ethnic group/race
- Global Region (North America, Eastern Europe, Western Europe and Australia, Asia, and Other)
- Menopausal status (postmenopausal, premenopausal)

### 9.3.2 Baseline Characteristics

#### **Baseline Characteristics and Disease History**

- Weight (kg)
- Height (cm)
- Body mass index (kg/m<sup>2</sup>)
- ECOG performance status at study entry (0, 1)
- Breast cancer therapy (neoadjuvant, adjuvant)
- Hormone receptor (ER/PR) status (positive, negative)
- HER-2 status (positive, negative)
- Molecular subtypes of breast cancer (hormone receptor positive and HER-2 positive, hormone receptor positive and HER-2 negative, hormone receptor negative and HER-2 positive, hormone receptor negative and HER-2 negative)
- Breast cancer AJCC stage (IA, IB, IIA, IIB, IIIA, IIIB, IIIC, IV)
- Primary tumor size (T-stage: T0, T1 or T1a-T1c or T1mic, T2, T3, T4 or T4a-T4d, Tx, Tis)
- Lymph node status (N-stage: N0 or pN0 or pN0 (i-) or pN0 (i+) or pN0 (mol-) or pN0 (mol+) or pN1mi, N1 or pN1 or pN1a - pN1c, N2 or N2a or N2b or pN2 or pN2a or pN2b, N3 or N3a - N3c or pN3 or pN3a - pN3a, Nx or pNx).
- Breast cancer histopathologic grade (G1=low, G2=intermediate, G3=high, not evaluable/missing)
- Time from initial breast cancer diagnosis to randomization (in months)

#### **Other Cancers Disease History**

- History of Other Cancers (yes, no)

#### **Substance Use History**

- Alcoholic beverages (none,  $\leq$ 2 per day,  $\geq$  3 per day)
- Tobacco use (never, former, current)

**Baseline Labs and Bone Turnover Markers (BTMs):**

- Calcium/ Albumin-adjusted calcium
- Magnesium
- Phosphorus
- Serum 25 (OH) Vitamin D
- Hematology
- Bone-specific alkaline phosphatase (BSAP)
- uNTx corrected for creatinine (uNTx/Cr)
- $\alpha$ CTx corrected for creatinine ( $\alpha$ CTx/Cr)

**Previous Anti-neoplastic Treatment and Surgical Procedure for Neoadjuvant / Adjuvant Subjects**

- Chemotherapy
- Anti-HER2 therapy
- Hormonal therapy
- Radiotherapy
- Breast cancer surgery

**9.3.3 Bone-specific Medication History**

- Subject incidence of historical bone-specific medication usage by preferred term

**9.3.4 Fracture History**

- Any historical fracture
- Historical osteoporotic fracture
- Historical vertebral fracture
- Historical non-vertebral fracture
- Parental hip fracture

**9.3.5 Baseline PRO Characteristics and Analgesic Use**

- Analgesic use
- BPI-SF
  - “Worst” pain score
  - Pain severity scale
  - Pain interference score
- EQ-5D
  - Health index scores
  - VAS

## 9.4 Safety Analyses

Safety analyses in this study will evaluate the safety profile of denosumab as compared with placebo. This will include assessments of treatment-emergent adverse events and adverse events of interest, investigational product exposure, concomitant medication usage, ECOG performance status, vital signs, immunogenic response, and clinical laboratory results. No statistical testing is planned in the safety analysis.

### 9.4.1 Adverse Events (AEs)

All adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 20.1 or higher. Treatment-emergent adverse events are events occurred during the time period from the first dose date (AEs are excluded if the indicator of “Did event start before the first dose of investigational product?” is checked as “Yes” on Adverse Events Summary CRF pages) through 30 days after the last dose of IP. Based on the AEs summary instructions, an AE that exists prior to the administration of IP and gets worse during the time period defined above will be included as a treatment emergent adverse event. Unless otherwise specified, all AE related analyses are based on treatment-emergent adverse events.

For all AEs, SAEs, grade  $\geq 3$  (ie, CTCAE grade 3, 4, and 5) AEs, fatal AEs, AEs leading to discontinuation of IP or study, as listed in [Section 4.2.1](#), subject incidence will be summarized by system organ class (SOC), high level term (HLT) and preferred term (PT), and by PT only. Subject incidence of AEs by SOC, HLT, PT, and worst grade will be tabulated. Subject-year adjusted incidence rates will be summarized for all AEs, SAEs, and grade  $\geq 3$  AEs by SOC, HLT and PT, and by PT only. A summary table for investigator determined IP related AEs will be provided by PT. In addition, subject incidence for all AEs and SAEs by SOC, HLT and PT will be tabulated by incremental time period (0 - 12 months, >12– 24 months, >24 - 36 months, >36 - 48 months, >48 months).

#### 9.4.1.1 Events of Interests (EOIs)

Events of interests (EOIs), including adverse events of hypocalcaemia, osteonecrosis of the jaw (ONJ), new primary malignancy, atypical femur fracture (AFF), cardiac/vascular disorders, infections and infestations, adverse events potentially associated with hypersensitivity, osteonecrosis excluding the jaw, musculoskeletal pain, and hypercalcemia occurred after 30 days following discontinuation of IP will be summarized. Subject incidence of AEs and SAEs of these EOIs will be tabulated at the EOI level and PT level.

#### **9.4.1.2 Osteonecrosis of the Jaw**

All subjects with an oral adverse event suspicious of osteonecrosis of the jaw (ONJ) should be examined by a dentist or other qualified oral specialist (eg, oral surgeon).

Potential events of ONJ are identified by: a) obtaining the available information from investigators on pre-specified oral event terms including those specifically reported as osteonecrosis of jaw; b) regular assessment of the clinical trial database to detect maxillofacial events, which might be indicative of ONJ; c) review and assessment of all these events by an independent adjudication committee. All adjudicated positive ONJ events with triggering adverse events having onset dates prior to the primary analysis cut-off date will be included. The proportion of subjects experiencing adjudicated positive ONJ will be summarized by PT. In addition, subject-year adjusted incidence rate of adjudicated positive ONJ by incremental time period (0 - 12 months, >12 – 24 months, >24-36 months, >36-48 months, >48 months) will be presented. A KM plot for time to first adjudicated positive ONJ will be provided.

#### **9.4.1.3 Atypical Femoral Fracture**

All subjects presenting with new or unusual thigh, hip, or groin pain should be evaluated for a suspected adverse event of atypical femoral fracture (AFF). Adverse events reported as AFF as well as adverse events identified by Amgen as potentially representing AFF will be reviewed by an independent adjudication panel of experts to determine whether the pre-defined criteria for AFF are met. Amgen will request the investigating site to provide all available source documents surrounding that event to be reviewed by the blinded adjudication committee. Adverse events that are adjudicated as positive for AFF and have onset dates prior to the primary analysis cut-off date will be included. The proportion of subjects experiencing adjudicated positive AFF will be summarized by PT. In addition, subject-year adjusted incidence rate of adjudicated positive AFF by incremental time period (0 - 12 months, >12 – 24 months, >24-36 months, >36-48 months, >48 months) will be presented. A KM plot for time to first adjudicated positive AFF will be provided.

#### **9.4.2 Investigational Product Exposure**

The number of months on study, number of doses received, and cumulative investigational product exposure will be summarized using descriptive statistics. FAS will be used for this summary.

#### **9.4.3 Concomitant Medications and Anti-neoplastic Therapies/Surgeries**

Subject incidence of bone-specific medications, and calcium and vitamin D medications will be summarized by medication preferred term. Safety analysis set will be used in the analysis.

On-study anti-neoplastic therapy (eg, chemotherapy, anti-HER-2 therapy, and hormonal therapy), radiotherapy and surgery for breast cancer will be summarized for subjects in the FAS. In case that there is a large imbalance on the use of on-study anti-neoplastic therapy between treatment groups, sensitivity analysis may be performed to adjust for the use for primary and key secondary endpoints.

#### **9.4.4 ECOG Performance Status**

The recorded value and change from baseline of ECOG performance status will be summarized by treatment group and visit using safety analysis set.

#### **9.4.5 Vital Signs**

Summary statistics for recorded values for systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse, respiration rate, and temperature will be displayed by visit using safety analysis set. Subjects with missing data for a clinical planned event will not contribute to the tabulations for that time point.

#### **9.4.6 Immunogenic Response**

Immunogenic response will be described by tabulating the numbers and percentages of subjects who tested positive (binding, neutralizing) for anti-denosumab antibodies in the subjects receiving at least one dose of denosumab. If a subject tests positive for neutralizing antibodies against denosumab, the relationship between the presence of neutralizing antibodies, adverse events, and bone turnover will be evaluated.

Immunogenic response will be listed by subject and visit date for subjects that test positive to either binding or neutralizing antibodies.

#### **9.4.7 Clinical Laboratory Results**

Only central lab parameters will be used for analysis. Laboratory parameters will be summarized over time via shift tables, in which the incidence of shifts of toxicity grade (CTCAE version 3.0 or higher) in recorded values from baseline to 'Worst' on-study value, are displayed by treatment group.

Laboratory parameters will also be summarized using descriptive statistics for recorded values and changes from baseline. Lab parameters will be summarized using

International System (SI) units. The following parameters: serum calcium adjusted for albumin, and phosphorus will also be summarized using conventional units.

In addition, the proportion of subjects who have grade shifts of at least 2 grades from baseline and the proportion of subjects who have CTCAE grade decrease of at least 2 grades in albumin-adjusted calcium at any time during the study will be summarized. The subject incidence of lab grade  $\geq 3$  (CTCAE grade 3 and 4) will also be tabulated.

Missing laboratory data are generally not imputed. If multiple laboratory records for a parameter are available on a day, the sample that is the closest to the scheduled time will be used. Laboratory data within a pre-specified window around a scheduled visit are considered valid. Pre-specified windows for each scheduled visit are supplied in the Appendix. In cases where multiple records exist within a window, the record that is the closest to the target day will be used. If 2 evaluations are the same distance from the target day, the later visit will be used.

## **9.5 Efficacy Analyses**

Efficacy analyses of the primary and secondary endpoints will determine the treatment effects of denosumab on bone metastasis-free survival (BMFS), disease-free survival (DFS) in the full study population, disease-free survival (DFS) in the postmenopausal subset, overall survival (OS), and distant recurrence-free survival (DRFS) compared with placebo.

The primary efficacy endpoint, BMFS, will be compared between denosumab and placebo using a log-rank test stratified by the randomization stratification factors. If denosumab is determined to be superior to placebo with respect to the primary efficacy endpoint, the secondary efficacy endpoints including

- Step 1: DFS in the full study population and DFS in the postmenopausal subset
- Step 2: OS
- Step 3: DRFS

will be tested in a stepwise fashion over 3 steps; ie, the treatment effect with respect to the secondary endpoints at the subsequent step will be tested only when denosumab is determined to be superior to placebo with respect to all the secondary endpoints at previous steps. At Step 1, DFS in the full study population and DFS in the postmenopausal subset will be tested simultaneously using Hochberg procedure to

adjust for multiplicity ([Sakamaki, 2013](#); Ye et al, 2011). The secondary endpoints will be analyzed using a similar stratified log-rank test.

In the interim / primary analysis, the primary and secondary endpoints will be tested in a stepwise fashion as described above with the critical p-values listed in [Table 11](#). In the interim analysis, the study could potentially be stopped if both BMFS and DFS in the full study population cross their corresponding efficacy boundaries, or if BMFS crosses the futility boundary. The other secondary endpoints OS and DRFS will only be tested hierarchically when denosumab is determined to be superior to placebo with respect to all three efficacy endpoints BMFS, DFS in the full study population and DFS in the postmenopausal subset in the interim /primary analysis. In order to strictly control the overall type I error rate, the final critical p-values will be updated based on the actual number of events achieved and the alpha levels that have been spent at the interim analysis ([Proschan et al, 2006](#)). If the tests on BMFS, DFS in the full study population and DFS in the postmenopausal subset are all successful at the primary analysis, similar procedure will be used to adjust for multiplicity between interim and primary analysis for OS and DRFS. Namely, the final critical p-values will be based on the actual numbers of events achieved at the interim and primary analysis, and the alpha levels that have been spent at the interim analysis using O'Brien Fleming spending function with information fraction based on actual numbers of events.

### **9.5.1 Primary and Secondary Efficacy Endpoints**

#### **9.5.1.1 Estimation of Outcomes**

The primary efficacy endpoint is BMFS, and the secondary efficacy endpoints are DFS in the full study population, DFS in the postmenopausal subset, OS, and DRFS.

#### **9.5.1.2 Hypotheses to Be Tested**

At the interim and primary analyses, the primary and secondary null hypotheses will be tested in a stepwise fashion (as described in [Section 9.5](#)), where the hypothesis of secondary endpoints will be tested only when the null hypothesis of BMFS is rejected at the designed significance level, and the secondary endpoints at the subsequent step will be tested only when the null hypothesis of all the secondary endpoints at previous steps is rejected.

- Primary null hypothesis: The overall risk of developing bone metastasis or death for the denosumab treated group is no less than that of the placebo treated group.
- Primary alternative hypothesis: The overall risk of developing bone metastasis or death for the denosumab treated group is less than that of the placebo treated



group. It is anticipated that the true hazard ratio of denosumab compared with placebo will be 0.8.

- Secondary null hypothesis for DFS in the full study population: The overall risk of disease recurrence or death for the denosumab treated group is no less than that of the placebo treated group.
- Secondary alternative hypothesis for DFS in the full study population: The overall risk of disease recurrence or death for the denosumab treated group is less than that of the placebo treated group. It is anticipated that the true hazard ratio of denosumab compared with placebo will be 0.83.
- Secondary null hypothesis for DFS in the postmenopausal subset: The risk of disease recurrence or death in the denosumab treated group is no less than that of the placebo treated group for subjects who have been postmenopausal at enrollment.
- Secondary alternative hypothesis for DFS in the postmenopausal subset: The risk of disease recurrence or death in the denosumab treated group is less than that of the placebo treated group for subjects who have been postmenopausal at enrollment. It is anticipated that the true hazard ratio of denosumab compared with placebo will be 0.76.
- Secondary null hypotheses for OS and DRFS:
  - The overall risk of all cause death for the denosumab treated group is no less than that of the placebo treated group.
  - The overall risk of distant recurrence or death for the denosumab treated group is no less than that of the placebo treated group.
- Secondary alternative hypotheses for OS and DRFS:
  - The overall risk of all cause death for the denosumab treated group is less than that of the placebo treated group.
  - The overall risk of distant recurrence or death for the denosumab treated group is less than that of the placebo treated group.

### 9.5.1.3 Primary Analysis

#### **Bone Metastasis-free Survival**

Kaplan-Meier estimates of the survival functions will be graphically displayed for each treatment group. Kaplan-Meier quartiles (25<sup>th</sup> percentile and median) with 2-sided 95% confidence intervals will be calculated if applicable. In addition, BMFS time will be summarized via displaying number of subjects at risk, the percent of subjects that are censored, and Kaplan-Meier event rates with 2-sided 95% confidence intervals at year 1, year 2, year 3, year 4 and year 5.

A log-rank test stratified by the randomization stratification factors (defined in [Section 3](#)) will be used to compare the BMFS of the two treatment groups. The hazard ratio of denosumab compared with placebo and its corresponding 2-sided 95% confidence interval will be estimated using a Cox proportional hazards model with treatment groups

as the independent variable and stratified by the randomization stratification factors. FAS will be used in the primary analysis.

**Disease-free Survival in the full study population, Disease-free Survival in the postmenopausal subset, Overall Survival and Distant Recurrence-free Survival**

If superiority of denosumab over placebo for BMFS is established, the key secondary endpoints of DFS in the full study population and DFS in postmenopausal subset will be tested simultaneously based on a Hochberg procedure. If denosumab is determined to be superior to placebo with respect to both of the key secondary efficacy endpoints, OS and DRFS will be tested hierarchically. If superiority is not demonstrated for an endpoint, analyses will be provided for subsequent endpoints (following the order as described above) using similar methods that are used for the primary endpoint. The analyses will be considered descriptive in nature. FAS will be used in the analysis.

**9.5.1.4 Supportive Analyses**

To demonstrate the robustness of the results for BMFS, DFS in the full study population, and DFS in the postmenopausal subset in the primary analysis, supportive analysis will be conducted in the following aspects:

- Analyses will be performed using the per protocol set.
- Analyses will be performed using the FAS and the actual strata.
- Analyses will be performed using the FAS without censoring subjects who had the first occurrence of event before randomization [ie, including baseline events]. For these subjects, they will have their event date default to their randomization date.
- Analysis will be performed for BMFS (based on the definition of bone metastasis [including investigator reported events]), DFS in the full study population (based on the definition of breast cancer disease recurrence [including investigator reported events]), and DFS in the postmenopausal subset (based on the definition of breast cancer disease recurrence [including investigator reported events]) using the FAS. For BMFS [including investigator reported events], subjects who have not experienced event will be censored at the later of the last onsite clinical visit date and the last assessment date (via central imaging analysis) for bone metastasis, or at the primary analysis data cut-off date, whichever comes first. Subjects who had first occurrence of bone metastasis [including investigator reported events] before randomization will be censored at their randomization date. For DFS in the full study population [including investigator reported events] and DFS in the postmenopausal subset [including investigator reported events], subjects who have not experienced event will be censored at the later of the last onsite clinical visit date and the last assessment date (via central imaging analysis) for disease recurrence, or at the primary analysis data cut-off date, whichever comes first. Subjects who had first disease recurrence [including investigator reported events] before randomization will be censored at their randomization date.

The following analysis will be performed for BMFS, DFS in the full study population, and DFS in postmenopausal subset using FAS to explore the potential informative censoring due to drop-out:

- Missing data subsequent to censoring will be imputed using the following two parameters:
  - The hazard rate subsequent to censoring in the control group is assumed to be varied from half to twice of the observed hazard rate before imputation
  - The hazard ratio for denosumab relative to control is assumed to be varied from half to twice of the observed hazard ratio before imputation.
  - The presentation of the ranges for assumed hazard rate in the control group and assumed hazard ratio may be reassessed based on the event distribution of observed data.
- For each combination of the two parameters:
  - Event status between censoring and primary data cut-off date for the control group will be randomly imputed using the hazard rate based on exponential distribution for BMFS and piece-wise exponential distribution for DFS. Similarly, event status between censoring and primary data cut-off date for the denosumab group will be randomly imputed using the hazard rate for the control group times the hazard ratio.
  - Repeat the imputation many times. The hazard ratio after imputation will be estimated for each repetition. The overall estimate of the hazard ratio will be calculated using the multiple imputation procedure ([Little and Rubin, 1987](#)).

#### 9.5.1.5 Subgroup Analyses

BMFS, DFS in the full study population, DFS in the postmenopausal subset will be analyzed within each of the subgroups indicated in [Section 6.3](#), except for age (<50 years, ≥ 50 years) and menopausal status (premenopause, postmenopause) for DFS in the postmenopausal subset.

For each category of a subgroup, the following analyses will be performed using the FAS:

- Kaplan-Meier quartiles (25<sup>th</sup> percentile and median) with 2-sided 95% confidence intervals, and Kaplan-Meier event rates with 2-sided 95% confidence intervals at year 1, year 2, year 3, year 4 and year 5 will be calculated if applicable.
- A log-rank test stratified by the randomization stratification factors (defined in [Section 3](#)) will be used to compare the two treatment groups. Hazard ratio of denosumab compared with placebo and its corresponding 2-sided confidence interval will be estimated using a Cox proportional hazards model with treatment groups as the independent variable and stratified by the randomization stratification factors.

- A Cox proportional hazards model with treatment groups, indicator variable of subgroups, the interaction between the treatment and indicator variable, as the independent variables, and stratified by the randomization stratification factors will be used. If the p-value of the interaction term is less than 0.05, then the interaction is significant, and the Gail and Simon test ([Gail and Simon, 1985](#)) will be used to test for qualitative interaction. The nature of the interaction will be investigated.

#### **9.5.1.6 Analysis of Covariates**

The covariates in [Section 4.4](#) will be analyzed for BMFS, DFS in the full study population, and DFS in the postmenopausal subset, except for age (<50 years,  $\geq 50$  years) and menopausal status (premenopause, postmenopause) for DFS in the postmenopausal subset. The 95% 2-sided confidence interval of the hazard ratio of denosumab compared with placebo, adjusted for each covariate separately and all covariates simultaneously from the Cox proportional hazards model will be obtained. FAS will be used. These covariates, except for the stratification variables, will be re-examined and appropriately re-categorized (due to small sample size) before unblinding for the interim and primary analyses.

#### **9.5.2 Exploratory Endpoints**

All analyses of exploratory efficacy endpoints listed in [Section 4.3](#) will be performed as follows:

##### **9.5.2.1 Time to Event Exploratory Endpoints**

For each of the time to event endpoint, ie, time to first bone metastasis (excluding death), time to first bone metastasis as site of first recurrence, time to disease recurrence, time to distant recurrence, time to first on-study fracture (vertebral or non-vertebral fracture), time to first on-study SRE, time to first on-study SRE or hypercalcemia, and time to on-study symptomatic bone metastasis,

- Kaplan-Meier method will be used to estimate the survival function by treatment group.
- A log-rank test stratified by the randomization stratification factors will be used to compare the two treatment groups.
- Hazard ratio of denosumab compared with placebo and its 2-sided 95% confidence interval will be estimated using a Cox proportional hazards model with treatment groups as the independent variable and stratified by the randomization stratification factors.
- In addition, for time to first bone metastasis as site of first recurrence and time to first symptomatic bone metastasis, cumulative incidence function (CIF) ([Kalbfleisch and Prentice, 2002](#)) will be estimated with non-bone site of first recurrence and with first asymptomatic bone metastasis as a competing event respectively. Comparison of the CIF between denosumab and placebo will be based on [Gray \(1988\)](#) and [Zhou et al \(2011\)](#) stratified by the randomization stratification factors.

FAS will be used in the analysis. It should be note that in the analysis of time to first bone metastasis as site of first recurrence and time to on-study symptomatic bone metastasis, subjects who developed the first recurrence outside the bone and subjects who developed bone metastases that were not asymptomatic are censored respectively. The censorship could be potentially informative, thus the results should be interpreted with caution.

#### **9.5.2.2 Analgesic Score**

Descriptive statistics for recorded values and changes from baseline will be displayed by visit. Proportion of subjects with AQA score  $\geq 3$  in subjects with baseline AQA score of 0-2 will be summarized at year 1, year 2, year 3, year 4 and year 5, and compared between two treatment groups using chi-square test at each year. FAS will be used in the analysis.

#### **9.5.2.3 PRO Analysis**

##### **BPI-SF**

For each of the BPI endpoint (“Worst” pain score, pain severity score and pain interference score) listed in [Section 4.3.2](#), descriptive statistics for recorded values and changes from baseline will be displayed by visit using the FAS.

The survival function of time to 1)  $\geq 2$ -point increase from baseline in “Worst” pain score, and 2)  $> 4$ -point in “Worst” pain score in subjects with baseline “Worst” pain score of 0-4 for each treatment group will be estimated using Kaplan-Meier method. The analyses will be based on the observed data. Analyses will include subjects who have an opportunity to meet the response criteria described above. For example, subjects with baseline score  $\leq 8$ -point will be considered in time to  $\geq 2$ -point increase analysis. The hazard ratio of denosumab compared with placebo and its corresponding 2-sided 95% CI will be estimated based on a Cox proportional hazards model adjusting the applicable covariates below and stratified by the randomization stratification factors (defined in [Section 3](#)).

- Treatment group (denosumab, placebo)
- Age (continuous)
- Race (Caucasian, non-Caucasian)
- Menopausal status (premenopause, postmenopause)
- Baseline ECOG=0 (yes, no)
- Time from breast cancer diagnosis to randomization
- Baseline pain at worst score

- Baseline analgesic score
- Breast cancer AJCC stage
- Global Region (North America, Eastern Europe, Western Europe and Australia, Asia, and Other)

The analysis described above will be repeated for the following two endpoints defined in [Section 5.7](#).

- Time to  $\geq 2$ -point increase score from baseline in 'Worst' pain score (after 6 months post randomization)
- Time to 4-point in "Worst" pain score (after 6 months post randomization)

The proportions of subjects who meet the following criteria for response will be summarized by visit.

- $\geq 2$ -point increase in 'Worst' pain score from baseline at each visit
- 'Worst' pain score  $> 4$ -point at each visit in subjects with baseline 'Worst' pain score of 0-4

The overall treatment effect will be tested based on all data and the data after 6 months post randomization respectively using Generalized Estimating Equations (GEEs) adjusting treatment, study visit, the interaction between treatment and study visit, baseline worst pain score, baseline analgesic score, IP dosing schedule (Q3W/Q3M IP dosing schedule, Q4W/Q3M IP dosing schedule) and randomized stratification factors. The analyses will be based on the observed data, and the imputed data using the PRO multiple imputation method. If the baseline score limits the response (eg, subject with baseline score  $\geq 9$ - point cannot reach  $\geq 2$ -point increase), subject is categorized as non-responder, and will be excluded from the analysis.

In addition, the proportions of subjects who meet the criteria for response above at year 1, year 2, year 3, year 4 and year 5 will be compared between the two treatment groups using a logistic regression model adjusting the same covariates listed above, IP dosing schedule (Q3W/Q3M IP dosing schedule, Q4W/Q3M IP dosing schedule), and randomization stratification factors. The odds ratio and the corresponding 95% 2-sided confidence interval and p-values will be provided. In case that the logistic regression model does not converge due to sparse data or separation, a conditional logistic regression stratified by randomization stratification factors may be used. The analyses will be based on the imputed data using the PRO multiple imputation method. If the baseline score limits the response (eg, subject with baseline score  $\geq 9$ -points cannot

reach a  $\geq 2$ -point increase), subject is categorized as non-responder, and will be excluded from the analysis.

#### **EQ-5D**

Descriptive statistics for recorded values and changes from baseline for the EQ-5D health index and EQ-5D VAS will be displayed by visit using the FAS.

#### **9.5.2.4 Breast Density**

The distribution of breast density level and the shift of breast density from baseline will be summarized by visit. The change from baseline in breast density will be summarized descriptively by visit. The comparison of the change in breast density from baseline between treatment groups at each visit will be based on a van Elteren stratified rank test ([van Elteren, 1960](#)) stratified by the randomization stratification factors.

#### **9.5.2.5 Pathological Complete Response**

For neo-adjuvant subjects, pathological complete response will be compared between treatment groups using a logistic regression model adjusting the randomization stratification factors.

#### **9.5.2.6 Residual Invasive Tumor Size**

For neo-adjuvant subjects, descriptive statistics for the residual invasive tumor size will be summarized, which includes n (number of non-missing observations), the mean, median, 25th and 75th percentiles, standard deviation, minimum, and maximum. The residual invasive tumor size between treatment groups will be compared using a van Elteren rank test ([van Elteren, 1960](#)) stratified by the randomization stratification factors. All neo-adjuvant subjects who are randomized and have observed values of residual invasive tumor size will be used in the analysis.

#### **9.5.2.7 Serum Denosumab Trough Concentrations**

Summary statistics will be calculated for serum denosumab trough concentrations. The PK analysis set defined in [Section 6.1.4](#) will be used in the analysis. For serum denosumab concentrations below the lower limit of quantification, a value of zero will be assigned for calculation of summary statistics.

#### **9.5.2.8 Blood Biomarkers and Bone Turnover Markers**

Blood biomarkers and bone turnover markers (BTMs; including uNTx/Cr,  $\alpha$ -CTx/Cr and BSAP), will be analyzed for all randomized subjects, as measured at baseline, and at each visit. Recorded values, change from baseline and percent change from baseline will be presented. Percent change in BTMs will be analyzed using non-parametric

methodology since it is anticipated that the distribution of these parameters will deviate from normality. A van Elteren stratified rank test ([van Elteren, 1960](#)) will be used to compare treatment groups with stratification factors. All the subjects who are randomized and have observed values of the endpoints at the time of interest will be used in the analysis.

#### **10. Final Analysis of Data From Ltfu**

Deaths during the long-term survival follow-up, and from randomization through LTFU will be summarized.



## 11. Literature Citations / References

Denosumab Investigator's Brochure. Thousand Oaks, CA: Amgen Inc.

Brooks R (1996) EuroQol: the current state of play. *Health Police*, 37(1): 53-72.

Cleeland CS. (1991) Research in cancer pain: what we know and what we need to know. *Cancer*, 67 (suppl 3): 823-7

Coleman, R., Cameron, D., Dodwell, D., Bell, R., Wilson, C., Rathbone, E., Keane, M., Gil, M., Burkinshaw, R., Grieve, R., Barrett-Lee, P., Ritchie, D., Liversedge, V., Hinsley, S., Marshall, H., on behalf of the AZURE investigators (2014). Adjuvant zoledronic acid in patients with early breast cancer: final efficacy analysis of the AZURE (BIG 01/04) randomised open-label phase 3 trial. *Lancet Oncol*, 15(9), 997-1006.

Colleoni, M., O'Neill, A., Goldhirsch, A., Gelber, R. D., Bonetti, M., Thurlimann, B., Price, K. N., Castiglione-Gertsch, M., Coates, A. S., Lindtner, J., Collins, J., Senn, H. J., Cavalli, F., Forbes, J., Gudgeon, A., Simoncini, E., Cortes-Funes, H., Veronesi, A., Fey, M., & Rudenstam, C. M. (2000). Identifying breast cancer patients at high risk for bone metastases. *J Clin Oncol*, 18(23), 3925-3935.

Cheng L, Swartz MD, Zhao H, et al (2012). Hazard of recurrence among women after primary breast cancer treatment - a 10-year follow-up using data from SEER-medicare. *Cancer Epidemiol Biomarkers Prev*. 21(5): 800-809.

Cox, D.R. (1972) Regression models and life tables (with discussion). *Journal of Royal Statistical Society, Series B*, 34: 187-220.

Fairclough, D.L. (2002) Design and Analysis of Quality of Life Studies in Clinical Trials Chapman & Hall/CRC.

Francis, P., Crown, J., Di Leo, A., Buyse, M., Balil, A., Andersson, M., Nordenskjold, B., Lang, I., Jakesz, R., Vorobiof, D., Gutierrez, J., van Hazel, G., Dolci, S., Jamin, S., Bendahmane, B., Gelber, R. D., Goldhirsch, A., Castiglione-Gertsch, M., & Piccart-Gebhart, M. (2008). Adjuvant chemotherapy with sequential or concurrent anthracycline and docetaxel: Breast International Group 02-98 randomized trial. *J Natl Cancer Inst*, 100(2), 121-133.

Gail, M. H., and Simon, R (1985) Testing for qualitative interactions between treatment effects and patients subsets. *Biometrics*, 41: 361-372.

Gray RJ (1988) A class of K-sample tests for comparing the cumulative incidence of a competing risk, *Annals of Statistics*, 16:1141-1154.

Howell, A., Cuzick, J., Baum, M., Buzdar, A., Dowsett, M., Forbes, J. F., Hochtin-Boes, G., Houghton, J., Locker, G. Y., & Tobias, J. S. (2005). Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. *Lancet*, 365(9453), 60-62.

Kalbfleisch, J. D., and R. L. Prentice (2002) The statistical analysis of failure time data (Second Edition). John Wiley & Sons, New York.

Kaplan, E.L., and Meier, P. (1958) Nonparametric estimation from incomplete observation. *Journal of the American Statistical Association*, 53: 457-481.

Lawless, J. F. (1982) Statistical Models and Methods for Lifetime Data, New York: John Wiley and Sons, Inc.

Little, R.J.A. and Rubin, D.B. (1987) Statistical Analysis with Missing Data. J. Wiley & Sons, New York.

Martin, M., Pienkowski, T., Mackey, J., Pawlicki, M., Guastalla, J. P., Weaver, C., Tomiak, E., Al-Tweigeri, T., Chap, L., Juhos, E., Guevin, R., Howell, A., Fornander, T., Hainsworth, J., Coleman, R., Vinholes, J., Modiano, M., Pinter, T., Tang, S. C., Colwell, B., Prady, C., Provencher, L., Walde, D., Rodriguez-Lescure, A., Hugh, J., Loret, C., Rupin, M., Blitz, S., Jacobs, P., Murawsky, M., Riva, A., & Vogel, C. (2005). Adjuvant docetaxel for node-positive breast cancer. *N Engl J Med*, 352(22), 2302-2313.

Proschan MA, Lan KKG, Wittes JT (2006). *Statistical Monitoring of Clinical Trials: A Unified Approach*. Springer, New York.

Romond, E. H., Perez, E. A., Bryant, J., Suman, V. J., Geyer, C. E., Jr., Davidson, N. E., Tan-Chiu, E., Martino, S., Paik, S., Kaufman, P. A., Swain, S. M., Pisansky, T. M., Fehrenbacher, L., Kutteh, L. A., Vogel, V. G., Visscher, D. W., Yothers, G., Jenkins, R. B., Brown, A. M., Dakhil, S. R., Mamounas, E. P., Lingle, W. L., Klein, P. M., Ingle, J. N., & Wolmark, N. (2005). Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med*, 353(16), 1673-1684.

Schoenfeld, D. (1982) Partial Residuals for the Proportional Hazards Regression Model. *Biometrika*, 69, 239 -241.

Sakamaki K (2013). A group sequential method using Hochberg procedure for clinical trials with multiple primary endpoints. Activity number 599, Presented at American Statistical Association Joint Statistical Meeting, Montreal, Canada, August 7, 2013.

Slamon, D., Eiermann, W., Robert, N., Pienkowski, T., Martin, M., Pawlicki, M., Chan, A., Smylie, M., Liu, M., Falkson, C., Pinter, T., Fornander, T., Shiftan, T., Valero, V., Mackey, J., Tabah-Fisch, I., Buyse, M., Lindsay, M., Riva, A., Bee, V., Pegram, M., Press, M., & Crown, J. (2006). BCIRG 006: 2nd interim analysis phase III randomized trial comparing doxorubicin and cyclophosphamide followed by docetaxel (ACT) with doxorubicin and cyclophosphamide followed by docetaxel and trastuzumab (ACTH) with docetaxel, carboplatin and trastuzumab (TCH) in Her2neu positive early breast cancer patients. *SABCS (Abstract #52)*.

Thurlimann, B., Keshaviah, A., Coates, A. S., Mouridsen, H., Mauriac, L., Forbes, J. F., Paridaens, R., Castiglione-Gertsch, M., Gelber, R. D., Rabaglio, M., Smith, I., Wardley, A., Price, K. N., & Goldhirsch, A. (2005). A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. *N Engl J Med*, 353(26), 2747-2757.

Van Elteren P.H. (1960). On the combination of independent two-sample test of Wilcoxon, *Bulletin of the International Statistical Institute*, 37, 351-361.

Ye Y., Li A., Liu L., Yao B. (2012). A group sequential Holm procedure with multiple primary endpoints. *Statistics in medicine*, 32: 1112-1124.

Zhou B, Latouche A, Rocha V, Fine J. (2011). Competing Risks Regression for Stratified Data. *Biometrics*. 67(2):661-70.

## **12. Appendices**

### **12.1 Technical Detail/Supplemental Information Regarding Statistical Procedures and Programs**

#### **12.1.1 Dates**

All dates should be converted to SAS dates prior to its use in any calculation.

#### **12.1.2 Visit Windows**

Per protocol, post baseline study visits are to be performed  $\pm 7$  days from the scheduled Q4W and Q3W visit dates, and  $\pm 14$  days from the scheduled Q3M visit dates. To allow for variations in scheduling, the following visit windows will be used to assign evaluations to a most appropriate nominal visit for analysis and summarization.

**For Subjects on the Q4W/Q3M Schedule:**

<b>Frequency of Assessments</b>	<b>Endpoint</b>	<b>Nominal Visit</b>	<b>Window Definition (Study Day)</b>
Q4W up to week 25 followed by Q12W up to week 253	Serum albumin, calcium and calcium (corrected)	Baseline Week 5 Week 9 Week 13 Week 17 Week 21 Week 25 Week 37 ... Week 241 Week 253	≤1 2 – 42 43 – 70 71 – 98 99 – 126 127 – 154 155 – 182 183 – 294 ... 1639 – 1722 ≥1723
Q12W up to week 25 followed by Q36W up to week 241	Magnesium, and phosphorus	Baseline Week 13 Week 25 Week 61 Week 97 Week 133 Week 169 Week 205 Week 241	≤1 2 – 126 127 – 210 211 – 546 547 – 798 799 – 1050 1051 – 1302 1303 – 1554 ≥1555
Multiple frequencies	Hematology, and vitamin D	Baseline EOTP	≤1 ≥2
Q12W	ECOG assessment, vital signs, and physical exam	Baseline Week 13 Week 25 Week 37 Week 49 ...	≤1 2 – 126 127 – 210 211 – 294 295 – 378 ...
Q12W up to week 277	PRO	Baseline Week 13 Week 25 Week 37 Week 49 ... Week 265 Week 277	≤1 2 – 126 127 – 210 211 – 294 295 – 378 ... 1807 – 1890 ≥1891
Q24W	Oral exam	Baseline Week 25 Week 49 Week 73 Week 97 ...	≤1 2 – 252 253 – 420 421 – 588 589 – 756 ...

<b>Frequency of Assessments</b>	<b>Endpoint</b>	<b>Nominal Visit</b>	<b>Window Definition (Study Day)</b>
Multiple frequencies	PK, BTMs and blood biomarkers	Baseline Week 13 Week 25 Week 37 Week 49 Week 61 Week 145	≤1 2 – 126 127 – 210 211 – 294 295 – 378 379 – 462 ≥463
Multiple frequencies	Anti-denosumab antibodies	Baseline Week 25 Week 49 Week 97 Week 145 Week 205 Week 253 Week 277	≤1 2 – 252 253 – 420 421 – 840 841 – 1176 1177 – 1638 1639 – 1932 ≥1933
Yearly	Breast density and PRO	Baseline Year 1 Year 2 Year 3 Year 4 Year 5	≤1 2 – 548 549 – 913 914 – 1278 1279 – 1643 ≥1644

**For Subjects on the Q3W/Q3M Schedule:**

<b>Frequency of Assessments</b>	<b>Endpoint</b>	<b>Nominal Visit</b>	<b>Window Definition (Study Day)</b>
Q3W up to week 19 followed by Q12W up to week 247	Serum albumin, calcium and calcium (corrected)	Baseline Week 4 Week 7 Week 10 Week 13 Week 16 Week 19 Week 31 ... Week 235 Week 247	≤1 2 – 31 32 – 52 53 – 73 74 – 94 95 – 115 116 – 136 137 – 252 ... 1597 – 1680 ≥1681
Q9W up to week 19 followed by Q36W up to week 235	Magnesium, and phosphorus	Baseline Week 10 Week 19 Week 55 Week 91 Week 127 Week 163 Week 199 Week 235	≤1 2 – 94 95 – 157 158 – 504 505 – 756 757 – 1008 1009 – 1260 1261 – 1512 ≥1513
Multiple frequencies	Hematology, and vitamin D	Baseline EOTP	≤1 ≥2
Q9W up to week 19 followed by Q12W	ECOG assessment, vital signs, and physical exam	Baseline Week 10 Week 19 Week 31 Week 43 ...	≤1 2 – 94 95 – 157 158 – 252 253 – 336 ...
Q9W up to week 19 followed by Q12W up to week 271	PRO	Baseline Week 10 Week 19 Week 31 Week 43 ... Week 259 Week 271	≤1 2 – 94 95 – 157 158 – 252 253 – 336 ... 1765 – 1848 ≥1849
Q18W up to week 19 followed by Q24W	Oral exam	Baseline Week 19 Week 43 Week 67 Week 91 ...	≤1 2 – 189 190 – 378 379 – 546 547 – 714 ...

Frequency of Assessments	Endpoint	Nominal Visit	Window Definition (Study Day)
Multiple frequencies	PK, BTMs and blood biomarkers	Baseline Week 10 Week 19 Week 31 Week 43 Week 55 Week 139	≤1 2 – 94 95 – 157 158 – 252 253 – 336 337 – 420 ≥421
Multiple frequencies	Anti-denosumab antibodies	Baseline Week 19 Week 43 Week 91 Week 139 Week 199 Week 247 Week 271	≤1 2 – 189 190 – 378 379 – 798 799 – 1134 1135 – 1596 1597 – 1890 ≥1891
Yearly	Breast density and PRO	Baseline Year 1 Year 2 Year 3 Year 4 Year 5	≤1 2 – 548 549 – 913 914 – 1278 1279 – 1643 ≥1644

For urine and serum samples, the baseline sample time must be **on or** before the 1<sup>st</sup> dose time. If there are multiple records within a Baseline (BL) window, the record that is the closest to and prior to and including Study Day 1 will be considered as the baseline value. If there are multiple records within the subsequent visit window, the record that is closest to the day of the planned visit (eg, target day for visit “Week 5” = 7 \* 5 – 7 + 1; target day for EOTP visit is the end of treatment phase date on the CRF) will be used as the measurement for this planned visit (the nominal visit). If two observations are the same distance from the target day, use the later one.

For samples that are assessed with different frequencies, a mixture of the windows given in Visit Table will be used. For example, if a parameter is evaluated every 4 week up to week 25 and then every 12 week thereafter, the window of Q4W will be used up to week 25. The window for week 37 will cover study day 183 to 294. From week 49 and onwards, the Q12W window will be used.

## 12.2 Reference Values / Toxicity Grades

The Common Terminology Criteria for Adverse Events (CTCAE) are available at the following link:

[http://ctep.cancer.gov/protocolDevelopment/electronic\\_applications/docs/ctcae3.pdf](http://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/ctcae3.pdf).

## 12.3 Patient Report Outcome Forms/Instruments

### 12.3.1 Patient Report Outcome Forms

In this study, the PRO measurements are: the BPI-SF and the EQ-5D.

### 12.3.2 Patient Report Outcome Instruments

#### 12.3.2.1 Reliability and Validity of PRO Instruments

The BPI-SF is a questionnaire specifically designed to assess pain in cancer. The BPI-SF captures information on the intensity of pain (pain severity) as well as the degree to which pain interferes with function (pain interference). The single item, which asks patients to rate their pain ‘at its worst’, is an exploratory endpoint in this study. The severity scale, which is comprised of 4 items, and the pain interference scale, which is comprised of 7 items, are also exploratory endpoints in this study. Evidence for reliability and validity of the BPI-SF has been well documented in cancer patients (Cleeland CS, 1991).

The EQ-5D, developed in 1990, is a widely used generic HRQOL instrument that allows for estimation of Quality Adjusted Life Years (QALY), a requirement for reimbursement dossiers and formulary submissions. The EQ-5D is comprised of 6 questions. The first 5 questions address the following quality of life dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety /depression. Each of these questions has a simple three-category response scale. The last question is represented by a 20-cm vertical VAS, scored from 0 to 100 asking the patient to “mark your own health state today” (Brooks, 1996).

#### 12.3.2.2 PRO Question Mapping

Question mapping for the PROs are outlined in the table below.

MEASURES (scale)	QUESTIONS
BPI-SF	
Pain Severity	3-6
Pain Interference	9A-9G
EQ-5D	
Mobility	8a
Self-Care	8b
Usual Activities	8c
Pain Discomfort	8d
Anxiety/Depression	8e
Vertical VAS	9



### **12.3.2.3 Scoring Algorithm for PROs**

#### **12.3.2.3.1 BPI-SF Scoring**

Raw scores (0-10) for single item question will be used for each item. The scoring for composite pain severity and pain interference base on the developer approved algorithm.

#### **12.3.2.3.2 EQ-5D Scoring**

The EQ-5D instrument yields two measures: health state index value and health state visual analog scale score. The scoring of each is detailed below.

##### **12.3.2.3.2.1 Health State Index Value**

Question 8, a-e, asks the subject to describe his/her present health state on five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each question has 3 response choices, listed in order of increasing severity. The first response choice (“no problems”) is coded as a 1; the second (“moderate problems”) is coded as a 2; the third (“extreme problems”) is coded as a 3, as indicated on the questionnaire. In addition to the five variables representing these items, a string variable formed by concatenating the values for the five dimension variables should be included in the database. For example, if a subject reports as, 2.a. = 2, 2.b. = 1, 2.c. = 1, 2.d = 1, and 2.e. = 2, then string variable “Health State” would be “21112”. In total, there are 243 health states.

Value classification for the 243 health states has been calculated in previous studies using regression models. Parameters estimated by the previous models can be applied to estimate a health state index value for this study using the dimension scores from question 8 a-e and the following set of IF / THEN statements:

#### **Calculation of Health State Index**

Generalized least-squares (GLS) regression techniques in which the functional form was additive were used to estimate health state index scores from the EQ-5D questionnaire. The dependent variable Y was defined as  $1 - S$  where “S” is the health state index score. Besides the intercept, the specification of the remaining independent variables was derived from the ordinal nature of items in the EQ-5D.

Since there are three response choices for each dimension, there are two dummy variables for each dimension. For example, there are two dummy variables for mobility (MO2, MO3), two dummy variables for self-care (SC2, SC3), two dummy variables for usual activities (UA2, UA3), two dummy variables for pain / discomfort (PD2, PD3), and

two dummy variables for anxiety/ depression (AD2, AD3). The table below shows the independent variables specified in the model. There is one dummy variable N3 represents whether any of the dimensions is at level 3. Dummy variable D23 represents whether any of the dimensions is not “1”. D23 variable will equal to 0 only for health state “11111” (when a patient marks “1” to all questions).

$$Y = \beta_1 D23 + \beta_2 MO2 + \beta_3 SC2 + \beta_4 UA2 + \beta_5 PD2 + \beta_6 AD2 + \beta_7 MO3 + \beta_8 SC3 + \beta_9 UA3 + \beta_{10} PD3 + \beta_{11} AD3 + \beta_{12} N3$$

EuroQoL dimension	Patient Response	Parameters
D23	D23=1 if a patient marks “2” or “3” to ANY question D23=0, otherwise (Note: D23=0 only for health state “11111”)	0.081
MO2	MO2= 1 if patient marks “2” to Mobility question MO2= 0, otherwise	0.069
SC2	SC2=1 if patient marks “2” to Self-Care question SC2=0, otherwise	0.104
UA2	UA2=1 if patient marks “2” to Usual Activities question UA2=0, otherwise	0.036
PD2	PD2=1 if patient marks “2” to Pain/Discomfort question PD2=0 otherwise	0.123
AD2	AD2=1 if patient marks “2” to Anxiety/Depression question AD2=0 otherwise	0.071
MO3	MO3=1 if patient marks “3” to Mobility question MO3=0 otherwise	0.314
SC3	SC3=1 if patient marks “3” to Self-Care question SC3=0 otherwise	0.214
UA3	UA3=1 if patient marks “3” to Usual Activities question UA3=0 otherwise	0.094
PD3	PD3=1 if patient marks “3” to Pain/Discomfort question PD3=0 otherwise	0.386
AD3	AD3=1 if patient marks “3” to Anxiety/Depression question	0.236
N3	N3=1 if patient marks “3” to any question N3=0 otherwise	0.269

The resulting scores can range between 1.0 and -0.594; scores less than 0.0 are included because a number of states were rated as worse than death. A higher score indicates a more preferred health status.

#### **12.3.2.3.2.2 Health State Visual Analog Scale**

Questionnaire Question 9 asks respondents to rate their present health status on a vertical 0 to 100 visual analog scale of 20 cm, with 0 labeled as “Worst imaginable health state” and 100 labeled as “Best imaginable health state.” The scale is marked in increments of “1,” with values labeled at each decile. The person scoring this question must observe the point at which the subject’s hand- drawn line intersects the scale and enter the closest integer.