CONFIDENTIAL

This document is provided for use by the editors of *Osteoporosis Int'l* and their designated peer-reviewers for review purposes only. It is not to be distributed further without permission from Merck.

If the manuscript associated with this protocol is accepted for publication, Merck will allow the journal to post on its website, at the time of publication, the key sections of the protocol that are relevant to evaluating the study, specifically those sections describing the study objectives and hypotheses, the patient inclusion and exclusion criteria, the study design and procedures, the efficacy and safety measures, the statistical analysis plan, and any amendments relating to those sections. Merck reserves the right to redact proprietary information.

Merck should be consulted before any elements of this protocol are publicly posted. Please contact:
boyd.scott@merck.com

018-04

A Phase III Randomized, Placebo-Controlled Clinical Trial to Assess the Safety and Efficacy of Odanacatib (MK-0822) to Reduce the Risk of Fracture in Osteoporotic Postmenopausal Women Treated With Vitamin D and Calcium

Product: MK-0822
Protocol/Amendment No.: 018-04

THIS PROTOCOL AND ALL OF THE INFORMATION RELATING TO IT ARE CONFIDENTIAL AND PROPRIETARY PROPERTY OF MERCK SHARP & DOHME CORP., A SUBSIDIARY OF MERCK & CO., INC., WHITEHOUSE STATION, NJ, U.S.A.

THIS PROTOCOL REPLACES THE ORIGINAL PROTOCOL AND ANY SUBSEQUENT AMENDMENTS AND SHOULD BE SIGNED BY ALL INVESTIGATORS SIGNING THE ORIGINAL PROTOCOL.

SPONSOR:

Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc. (hereafter referred to as the **SPONSOR** or **Merck**)
One Merck Drive
P.O. Box 100
Whitehouse Station, NJ 08889-0100, U.S.A.

Protocol-specific Sponsor Contact information can be found in the Administrative Binder.

TITLE:

A Phase III Randomized, Placebo-Controlled Clinical Trial to Assess the Safety and Efficacy of Odanacatib (MK-0822) to Reduce the Risk of Fracture in Osteoporotic Postmenopausal Women Treated With Vitamin D and Calcium

INVESTIGATOR:

PRIMARY:

CLINICAL PHASE: III

US IND NUMBER: 70,893

SITE:

INSTITUTIONAL REVIEW BOARD/ETHICS REVIEW COMMITTEE:

Product: MK-0822 1
Protocol/Amendment No.: 018-04

SUMMARY OF CHANGES

PRIMARY REASON FOR THIS AMENDMENT:

This amendment restores the AST/ALT/bilirubin discontinuation criteria, to those previously used in MK-0822-018-02, following a regulatory agency request. A definition of Events of Clinical Interest (ECI) based on these laboratory criteria is also provided.

All specific changes are outlined below.

OTHER CHANGES INCLUDED IN THE AMENDMENT:

Section	Revision
Table of Contents	Corrected section header numbering inconsistencies noted in the table of contents and body of text for both Word and PDF versions of the amendment.
2.5, 2.7	Added the following to sub-section Height/Stature: "Stadiometers must be calibrated according to pre-specified procedures. Height measurements may be excluded if obtained with stadiometers that have not been adequately calibrated." "Height measurements obtained with stadiometers that have not been adequately calibrated will be excluded from the analysis."
3.1.1.1	Updated Risks sub-section with more recent data: "As of January 2012, odanacatib has been studied in approximately 576 healthy male and female subjects enrolled in 23 phase I studies" "As of December 2010, 129 patients completed Year 5. An open-label extension of this study is ongoing, with a total treatment duration of 10 years being planned. As of December 2011, 115 study participants completed Study Year 6"
3.1.6.	Removed the following text on genetic sample collection: "Ultimately, it will be necessary to validate the findings (if any) from this study, and this might possibly be accomplished through the *ad hoc* division of the study set, whereby one part is used as a training subset, and the other part as a testing subset. If this is not possible, for example due to low sample numbers and statistical power, then this study will be used solely as a training set, and subsequent and independent clinical studies will provide for the testing and validation of its scientific findings."

Product: MK-0822 2
Protocol/Amendment No.: 018-04

Section	Revision
3.2.1	Updated the anti-fungals subsection with more recent data: "Subsequent preclinical studies were conducted where the corresponding safety margins to clinical exposure achieved were ~9-fold in skeletally mature monkeys and ~12-fold in skeletally mature dogs for bone findings, while the preclinical safety margin for soft-tissues was about 12-fold in skeletally immature monkeys. These margins are far higher than the 2.4-fold increase in exposure that could be potentially resulted from concomitant treatment of odanacatib with a strong CYP3A4 inhibitors such as ketoconazole"
3.2.3.1.2	Updated section on study conduct for genetic sample collection. Added: "Subjects whose previously collected genetic samples have been inadvertently compromised (e.g. sample integrity, mislabeling, etc) may be invited to re-consent and provide a replacement genetic sample to be included in the genetic analyses. This will be done to ensure that all populations studied in this trial are properly represented." Removed: "At the time of sample collection, a witness from the investigator's staff should verify that the subject/patient has signed the consent and the correct subject/patient-specific label is placed on the genetic sample."
3.2.3.6.7	Removed the following foot note from Table 3-7 (Month 24, 36 and 48 procedures) as the trial is event-driven and will continue beyond Month 48 for a number of patients: "Study supplies will not be dispensed or diary cards provided at Month 48 or final study visit (if this is prior to Month 48)."
3.2.3.8	Clarified the following statement with regard to emergency unblinding: "Every effort should be made to contact the Clinical Monitor prior to such unblinding, *however the Investigator may unblind a patient for safety reasons without first contacting the Clinical Monitor.*" "Note that all patients who have been unblinded must be discontinued from the study"
3.2.3.9.1	Restored ALT/AST/bilirubin laboratory discontinuation criteria to those previously outlined in MK-0822-018-02. The restored discontinuation guidance, consistent with that in P018-02, now mandates the following for discontinuation of blinded study therapy: "Persistent elevations [> 3 x Upper Limit of Normal (ULN)] in aspartate aminotransferase (AST) or alanine aminotransferase (ALT) (2 consecutive readings at least 2 weeks apart) - OR - Persistent elevations [> 2 x Upper Limit of Normal (ULN)] in total bilirubin (2 consecutive readings at least 2 weeks apart)."

Section	Revision
3.3.1.3	Updated section to add the following: "Stadiometers must be calibrated according to pre-specified procedures. Height measurements obtained with stadiometers that have not been adequately calibrated will be excluded from the statistical analysis."
3.3.4	Updated adjudication procedures (Fractures and Delayed Fracture Union AEs sub-section) with the definition of osteoporotic and traumatic fractures: "As is the case in all fracture endpoint trials, a determination will be made for each incident clinical fracture as to whether it is **osteoporotic** (defined as fractures that occur in the absence of trauma or in a low impact trauma setting that would not have resulted in fracture in an individual *without* osteoporosis), **traumatic** (i.e., secondary to excessive force capable of causing a fracture in an individual *without* osteoporosis) or due to another cause (e.g, tumor or stress fracture from repetitive low energy force"
3.4.5.2	Updated definition of a Event of Clinical Interest (ECI) to: "Persistent elevations [> 3 x Upper Limit of Normal (ULN)] in aspartate aminotransferase (AST) or alanine aminotransferase (ALT) (2 consecutive readings at least 2 weeks apart) - OR - Persistent elevations [> 2 x Upper Limit of Normal (ULN)] in total bilirubin (2 consecutive readings at least 2 weeks apart)." Of note, this definition is aligned with discontinuation rule outlined in Section 3.2.3.9.1.
3.5.5.1	In addition to the planned efficacy analyses for non-vertebral, hip and clinical vertebral fractures, the following statement was added: "As a sensitivity analysis, similar analyses will be performed for all adjudicated clinical hip fractures, irrespective if they were osteoporotic (defined as fractures that occur in the absence of trauma or in a low impact trauma setting that would not have resulted in fracture in an individual *without* osteoporosis), traumatic (i.e., secondary to excessive force capable of causing a fracture in an individual *without* osteoporosis), stress or pathological. A similar sensitivity analysis will also be performed for all adjudicated non-vertebral, adjudicated vertebral and all adjudicated fractures."
Appendix 6.7	The protocol violation criterion, "Baseline 25-hydroxyvitamin D level <9 ng/mL" was removed since this test was not assessed in all patients and was not a requirement for study entry.

PROTOCOL

A Phase III Randomized, Placebo-Controlled Clinical Trial to Assess the Safety and Efficacy of Odanacatib (MK-0822) to Reduce the Risk of Fracture in Osteoporotic Postmenopausal Women Treated With Vitamin D and Calcium

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT.)

TABLE OF CONTENTS (CONT.)

TABLE OF CONTENTS (CONT.)

TABLE OF CONTENTS (CONT.)

TABLE OF CONTENTS (CONT.)

1. SUMMARY

1.1 TITLE

A Phase III Randomized, Placebo-Controlled Clinical Trial to Assess the Safety and Efficacy of Odanacatib (MK-0822) to Reduce the Risk of Fracture in Osteoporotic Postmenopausal Women Treated With Vitamin D and Calcium

1.2 INDICATION

Treatment of osteoporosis in postmenopausal women to prevent fractures.

1.3 SUMMARY OF RATIONALE

Osteoporosis currently affects approximately 200 million adults worldwide [1]. Approximately 30% of all postmenopausal women in the United States and in Europe have osteoporosis [2]. At least 40% of these women will sustain one or more fragility fractures of the hip, vertebrae, wrist or ribs in their remaining lifetimes. According to data available in 2002, the total annual direct costs for health care attributable to osteoporotic hip fractures was estimated at $18 billion in the US, $5.7 billion in Japan, and $4 billion in the EU [3; 4]. As such, osteoporosis poses a considerable burden both on the health care system and on society at large. The number of patients who will benefit from osteoporosis therapy is growing steadily and is expected to continue to do so for the foreseeable future, since fewer than 25% of osteoporotic patients receive treatment for osteoporosis, the society is aging, and osteoporosis in men is increasingly recognized as a medical issue.

Current treatment options for osteoporosis include bisphosphonates such as alendronate, risedronate and ibandronate, estrogens such as Premarin, selective estrogen receptor modulator (SERMs) such as raloxifene and its analogues, parathyroid hormone, calcitonin, strontium and supplements such as calcium and vitamin D_3. Other than alendronate, zoledronic acid, and hormone replacement therapy, currently available treatments for osteoporosis have demonstrated fracture reduction efficacy only at vertebral and non-vertebral sites. Safety and tolerability limitations exist for most osteoporotic agents. For the bisphosphonates, these include upper GI toxicity and irritation, renal toxicity, and hypothetical concerns about long-term skeletal residence. Osteonecrosis of the jaw in bisphosphonate users has been reported in the published literature. Long term use of estrogens has been shown to increase the risk of myocardial infarction, stroke, fluid retention, deep vein thrombosis and breast cancer. Parathyroid hormone causes concerns because of the observation of osteosarcomas in rats. Consequently there is an unmet medical need for osteoporosis treatments which are highly efficacious but have improved safety and tolerability profiles compared to the currently available agents.

Osteoclastic bone resorption requires demineralization of inorganic bone mineral followed by degradation of organic bone matrix. These processes occur sequentially via two separate mechanisms. The first process involves the secretion of acid into resorption

lacunae on the bone surface; the second consists of the degradation of organic (mainly proteinaceous) matrix by cathepsin K. Cathepsin K, which is abundantly expressed in osteoclasts, is a cysteine protease which exhibits collagenolytic activity under acidic conditions. Confirmation of the skeletal activity of cathepsin K in humans comes from the rare hereditary bone disorder, pycnodysostosis, which results from the presence of a defective cathepsin K gene and which is associated with an osteopetrotic phenotype.

Odanacatib is a potent, orally-active inhibitor of cathepsin K which is being developed for the treatment of postmenopausal osteoporosis. Odanacatib has demonstrated robust efficacy in preclinical models, with dose-dependent increases in bone mineral density (BMD) in ovariectomized rabbits and suppression of urinary N-telopeptide cross-links (uNTx) in ovariectomized (OVX) rhesus monkeys. The combination of increased BMD and suppression of bone resorption markers predicts fracture risk reduction in the clinic. In contrast to currently available anti-resorptive osteoporosis therapies cathepsin K inhibitors are expected to exhibit only limited suppression of bone formation. This hypothesis is supported by observations of normal bone formation in cathepsin K null mice and by preservation of bone formation in ovariectomized rabbits treated with a cathepsin K inhibitor for 6 months. It has also been reported that biochemical markers of bone formation, including osteocalcin and bone specific alkaline phosphatase, were not suppressed by the cathepsin K inhibitor balicatib in a 12 month Phase IIb study [5]. The absence of suppression of bone formation in tandem with inhibition of bone resorption (a phenomenon referred to as "reduced suppression of bone formation") may provide an advantage over existing anti-resorptive therapies, all of which suppress both resorption and formation.

Odanacatib has several distinguishing characteristics that may confer safety advantages over other cathepsin K inhibitors. Odanacatib is a non-basic molecule and does not accumulate in the acidic cellular compartment of the lysosome; it is also more selective for cathepsin K than for cathepsins B (>200-fold) and L (>1000-fold). This selectivity for cathepsin K is demonstrated in both assays using isolated enzymes and in whole cell assays (which contain lysosomes). By contrast, balicatib retains only 5-6 fold selectivity for cathepsin K over cathepsins B and L in whole cell assays even though it is more selective for cathepsin K in assays using the isolated enzymes. Since cathepsin B has a wide tissue distribution (including the skin) and is involved in apoptosis and collagen turnover, and cathepsin L is mainly involved in epidermal homeostasis, off-target activity of balicatib may explain the cutaneous adverse event profile which has been observed in clinical trials with this drug. Relacatib is non-specific and is equally selective for cathepsins K, L and V. Consequently, off-target activity of relacatib is likely to play a role in the adverse event profile observed. As such, drug-related adverse events seen with balicatib and relacatib are not predicted to arise with odanacatib due to their very different specificity profiles.

Overall, it is anticipated that odanacatib will demonstrate efficacy which is at least similar to that of the bisphosphonates, but without any risk for esophageal irritation. Furthermore, given the rapidity (compared to bisphosphonates) with which odanacatib is cleared from the bone, there is potential for the treatment of younger adult patients (with

anorexia nervosa or hypothalamic amenorrhea, for example) whose physicians currently prescribe bisphosphonates with some reluctance.

1.4 SUMMARY OF STUDY DESIGN

Study Design This is an event-driven, double-blind, randomized, placebo-controlled, multi-center, worldwide study. This trial will enroll osteoporotic postmenopausal women who have either no prior vertebral fractures or one prior vertebral fracture. Odanacatib has not been associated with an increase in skin or respiratory adverse experiences in clinical or preclinical studies. However, because the adverse experiences of morphea-like skin lesions and upper respiratory tract infections have been reported in a phase IIb study of the non-Merck cathepsin K inhibitor balicatib [5], a regulatory agency has requested that the SPONSOR enroll the current trial in two separate phases. This two-phase approach to study enrollment is designed to avoid unnecessarily exposing large numbers of patients to odanacatib, and to permit study of a limited number of patients with close monitoring by an independent Data Monitoring Committee (DMC) to demonstrate that adverse experiences similar to those seen with balicatib are not associated with odanacatib treatment. In the first phase of enrollment, approximately 1500 patients were randomized in a 1:1 ratio to receive vitamin D_3 (5600 IU once weekly), alone or in combination with odanacatib (50 mg once weekly). Patients also received a sufficient supply of open-label daily calcium supplements, supplied as calcium carbonate, so that their total daily calcium intake (from both dietary and supplemental sources) was approximately 1200 mg. After approximately 1500 patients were enrolled, enrollment was interrupted until all patients in the 'Lead Cohort' received study drug or placebo for at least 9 months. During this 9-month period, data were reviewed by the DMC approximately every 4 months, corresponding to an incremental increase in study drug exposure of approximately 250 patient-years between DMC reviews. After the 9-month safety data were analyzed and reviewed by the DMC as outlined in Section 3.5.5.6, and the risk/benefit ratio was found to be reassuring, the second phase of recruitment began enrolling the balance of ~16,300 patients, the 'Main Cohort'. The primary endpoints in this study are the cumulative incidence of vertebral, non-vertebral and hip fractures which are ordered, as described in Section 3.5.5.1, as follows: First, morphometric vertebral fractures, then second, non-vertebral and hip fractures. The trial is fracture event-driven, and it will end when the pre-specified number of patients with an incident fracture event required for efficacy demonstration has been reached. To demonstrate fracture risk reduction (assuming efficacy similar to that of alendronate) at the spine, hip and non-vertebral sites, the number of patients required to have an incident fracture event are 114, 237, and 824, respectively (Tables 3-13 and 3-14 in Section 3.5.5.4). Patients who discontinue blinded study therapy must continue to be followed in the study per-protocol. However, note that patients inappropriately randomized into the trial (i.e. those who do not meet entry criteria) may be discontinued from the study at the discretion of the SPONSOR.

Sample Size The precise sample size for the study has not been pre-specified, but is expected to lie between 12,000 and 20,000 study participants depending on the ratio of patients with a prior vertebral fracture to those without a prior vertebral fracture who

enroll in the study. According to data obtained from participating investigational sites 'Lead Cohort', the anticipated ratio of patients with a prior vertebral fracture to those without a prior vertebral fracture is approximately 1:3. Based on the overall fracture risk of this mixed population, the sample size is predicted to be approximately 16,300, comprised of approximately 4100 patients with a prior vertebral fracture and approximately 12,200 patients without a prior vertebral fracture. The size of the sample may be revised downward if a higher than anticipated proportion of patients with a prior fracture is enrolled. Conversely, it may be revised upward if a lower than expected proportion of patients with prior fracture is enrolled, or if additional exposure to assess safety is desired.

Trial Duration The study is expected to be approximately 5 years in duration, including 1 year for enrollment, and is not anticipated to exceed a total of 4 years duration for study participants in the 'Main Cohort'. For those in the 'Lead Cohort', trial duration could be approximately 5 years. Interim analyses will be performed after approximately 70% and 85% of hip fracture events have occurred so as to retain the possibility of early trial termination for efficacy (see Data Analysis Section 3.5.5.6). Patients will be monitored for excessive bone loss throughout the study. Those patients determined to have excessive bone loss will be discontinued from the blinded study therapy for treatment with available therapy.

1.5 SAMPLE

The patients included in the study will be representative of the general population of postmenopausal women with osteoporosis. As described in the preceding section, the sample size is expected to lie between 12,000 and 20,000 patients. The trial will enroll postmenopausal women who are ≥ 65 years old. Patients without a prior vertebral fracture must have a total hip or femoral neck bone mineral density T-score ≤ -2.5. Patients with a prior vertebral fracture may have a total hip or femoral neck bone mineral density T-score ≤ -1.5. Patients in the 'no prior vertebral fracture' group will be enrolled such that at least 2/3 will be ≥ 70 years old. Patients with more than one pre-existing vertebral fracture, or bone mineral density T-score <-4.0 at the total hip or femoral neck will not be eligible unless they are unsuitable for or decline osteoporosis therapy proven effective at non-vertebral sites (i.e., bisphosphonates, strontium, or PTH). Patients may not have had a prior hip fracture, a clinical (symptomatic) vertebral fracture within the past 24 months, and may not have a concomitant illness or laboratory abnormality which might preclude trial completion or confound data interpretation of study results (e.g., history of metabolic bone disease other than osteoporosis). Treatment with bone-active agents (e.g., bisphosphonates, estrogens, glucocorticoids) is limited before and during the trial as specified in Section 2.3 and Section 3.2.1.

1.6 DOSAGE/DOSAGE FORM, ROUTE, AND DOSE REGIMEN

Patients will receive either odanacatib (50 mg once weekly) or placebo. All patients will receive vitamin D_3 (5600 IU once weekly). Patients will also receive a sufficient supply of open-label daily calcium supplements, supplied as calcium carbonate, so that their total daily calcium intake (from both dietary and supplemental sources) is approximately 1200 mg.

1.7 STUDY FLOW CHART

Visit Number	Base Study																			
	1	2	3	4	5	6	7T	8	9T	10	11T	12	13T	14	15 T	16	17 T	18		
Study Time Point	Day 0 Screen	Day 1 Rand	Mo 3	Mo. 6.	Mo. 9	Mo. 12	Mo. 15	Mo. 18	Mo. 21	Mo. 24	Mo. 27	Mo. 30	Mo. 33	Mo. 36[11,12]	Mo. 39	Mo. 42	Mo. 45	Mo. 48[13]	EOS	14 Day F/U[14]
Visit Window	N/A	≤ 1 Mo. (30 days)	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.		
Procedures																				
Obtain informed consent for study	x																			
Review inclusion/exclusion criteria	x	x																		
Collect medical history	x																			
Review prior medication	x																			
Perform BMD measurement																				
Total hip[1] & subregions	x[4]			x[10]		x				x				x				x	x	
lumbar spine	x[9]	x[9]		x[10]		x				x				x				x	x	
distal forearm[2]		x				x				x				x				x	x	
total body[2]		x				x				x				x				x	x	
Perform limited physical exam (including vital signs and weight; height at 12 months and annually thereafter)	x		x	x	x	x		x		x		x		x		x		x		
Perform complete physical exam (including vital signs, height, and weight)[3]		x												x[15]					x[15]	
Perform lateral spine x-ray[3,4]	x			x		x				x				x				x	x	
Collect blood and urine lab safety assessments[3]	x		x	x	x	x		x		x		x		x		x		x	x	
Collect blood and urine for lab efficacy assessments and archives[2,5]		x		x		x				x				x				x	x	
Meal questionnaire in first 1500 patients enrolled			x	x	x															
PK sample in first 1500 patients enrolled		x	x	x	x															
Perform ECG		x																		
Calcium Questionnaire	x					x				x				x				x	x	
Dispense study medication	x[8]	x	x	x	x	x		x		x		x		x		x				
Provide diary card[6]	x	x	x	x	x	x		x		x		x		x		x				
Review diary card		x	x	x	x	x		x		x		x		x		x		x	x	
Perform tablet count for compliance[7]			x	x	x	x		x		x		x		x		x		x	x	
Review adverse experiences		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Review concomitant medication		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	

Visit Number	Base Study																			
	1	2	3	4	5	6	7T	8	9T	10	11T	12	13T	14	15 T	16	17 T	18		
Study Time Point	Day 0 Screen	Day 1 Rand	Mo. 3	Mo. 6	Mo. 9	Mo. 12	Mo. 15	Mo. 18	Mo. 21	Mo. 24	Mo. 27	Mo. 30	Mo. 33	Mo. 36[11,12]	Mo. 39	Mo. 42	Mo. 45	Mo. 48[13]	EOS	14 Day F/U[14]
Visit Window	N/A	≤ 1 Mo. (30 days)	±3 wks	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.	±3 wks.		
Perform Health Resource Utilization (if fracture has occurred)		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
Plasma for proteomics archive		x				x								X[17]					X[17]	
Obtain informed consent for genetic sampling		May be performed at any post-screening clinic visit																		
Collect genetic sample																				
Perform x-rays for cause at any visit																				
Obtain informed consent for bone biopsy[16]										M23				M35						
Dispense bone biopsy labeling agent[16]										M23				M35						
Perform bone biopsy[16]										x				x						

[1] Use left hip at screening, unless left hip is not evaluable. For subsequent hip scans, perform scan on the same hip as was scanned at baseline.
[2] Distal forearm, and total body BMD done in a randomly selected 10% of patients at Randomization, Mo. 12, then annually thereafter; efficacy labs (biochemical markers of bone turnover, and PTH) done in a randomly selected 10% of patients at Randomization, Mos. 6, 12, then annually thereafter; 25-hydroxyvitamin D measured at Randomization, Mos. 12 and end of study; this is the same 10% subset for all measurements.
[3] May be performed at other visits for cause.
[4] Perform BMD only if patient already qualifies by medical history and medication review; perform baseline lateral spine x-ray only if patient qualifies by preliminary (site) review of screening BMD.
[5] Serum and urine archives obtained at each time point indicated in Study Flow Chart in the 10% subgroup. Serum and urine archives to be obtained in *all* patients at Randomization, Month 12, and end of study.
[6] Dosing with blinded study therapy and study-provided Vitamin D_3 should be recorded on diary cards.
[7] Study medication only.
[8] Only open-label vitamin D_3 and calcium at Visit 1. Dispense these only to potentially eligible patients (e.g., based on total hip/femoral neck BMD and other parameters evaluable at the screening visit [e.g. medical history, medication review, etc.]).
[9] Spine DXA at screening will *only* be performed in cases where this is required by regulatory agency via documented request, subsequently approved by the SPONSOR. Patients who had BMD at the lumbar spine performed at screening (Visit 1) do not need to perform this procedure at Randomization (Visit 2).
[10] Spine/hip DXA at 6 months will *only* be performed in cases where this is required by regulatory agency via documented request, subsequently approved by the SPONSOR. As with all follow-up DXA measurements, results will be masked unless Excessive Bone Loss (EBL) criteria are met.
[11] Perform at Month 36 or end of study/early discontinuation if before Month 36.
[12] Given that this is an event-driven study for morphometric vertebral, hip, and non-vertebral fractures, actual study duration is unknown, but is estimated to be 5 years including a 1 year recruitment period.
[13] For patients with visits past 48 months, Telephone visits will be conducted every three months (e.g., Months 51, 57) procedures same as Month 45; a clinic visit for safety parameters every 6 months (e.g., Months 54, 66) procedures same as Month 42;, and complete visit annually (e.g., Month 60) procedures same as 48. See section 3.2.3.6.9.
[14] Telephone contact at least 14 days after the last dose of blinded study therapy or study discontinuation/end of study, whichever occurs later to collect SAE information.
[15] Complete PE performed at end of study visit.
[16] Bone biopsy at Month 24 and/or Month 36.
[17] Plasma for proteomics archive to be obtained in all patients at Randomization, Month 12, and end-of-study (may not be Month 36).

T=Telephone follow-up; Mo.=Month; IRB=Institutional Review Board; ERC=Ethics Review Committee; F/U=Follow-up; EBL=Excessive Bone Loss

2. CORE PROTOCOL

2.1 OBJECTIVES AND HYPOTHESES

In postmenopausal women with osteoporosis:

2.1.1 Primary

(1) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on the risk of morphometrically assessed vertebral fractures compared to placebo.

Hypothesis: Treatment with odanacatib reduces the risk of morphometrically assessed vertebral fractures compared to placebo.

(2) **Objective:** To assess the effect of treatment with odanacatib on the risk of hip fractures compared to placebo.

Hypothesis: Treatment with odanacatib reduces the risk of hip fractures compared to placebo.

(3) **Objective:** To assess the effect of treatment with odanacatib on the risk of clinical non-vertebral fractures compared to placebo. (Note: for the purposes of this study, non-vertebral fractures exclude fractures of the fingers, toes, face, and skull.)

Hypothesis: Treatment with odanacatib reduces the risk of clinical non-vertebral fractures compared to placebo.

2.1.2 Secondary

(1) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on the risk of clinical vertebral fractures compared to placebo.

Hypothesis: Treatment with odanacatib reduces the risk of clinical vertebral fractures compared to placebo.

(2) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on height compared to placebo.

Hypothesis: Treatment with odanacatib reduces height loss compared to placebo.

(3) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on the lumbar spine, total hip, femoral neck, trochanter and distal forearm BMD compared to placebo.

Hypothesis: Treatment with odanacatib increases lumbar spine, total hip, femoral neck, trochanter, and distal forearm BMD compared to placebo.

(4) **Objective:** To assess the safety and tolerability of treatment with odanacatib 50 mg once weekly compared to placebo.

Hypothesis: Odanacatib is safe and well tolerated compared to placebo.

(5) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on biochemical indices of bone resorption (serum C-Telopeptides of Type 1 collagen [s-CTx] and urine N-Telopeptides of Type 1 collagen [u-NTx]) compared to placebo.

Hypothesis: Treatment with odanacatib decreases biochemical indices of bone resorption (s-CTx and u-NTx) compared to placebo.

(6) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on lumbar spine, total hip, femoral neck, and hip trochanter BMD compared to placebo, in patients who are bisphosphonate-intolerant (i.e. patients with a contraindication or history of intolerance to bisphosphonates, or those considered by their physician to be unsuitable for bisphosphonate treatment).

Hypothesis: Treatment with odanacatib increases lumbar spine, total hip, femoral neck, and hip trochanter BMD compared to placebo in patients who are bisphosphonate-intolerant (i.e. patients with a contraindication to or a history of intolerance of oral bisphosphonate use).

(7) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on biochemical indices of bone formation (serum bone-specific alkaline phosphatase [BSAP] and serum N-Terminal Propeptides of Type 1 collagen [s-P1NP]) compared to placebo.

(8) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on qualitative histomorphometry of transilial bone biopsy specimens compared to placebo.

2.1.3 Exploratory Objectives

(1) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on total body BMD compared to placebo.

(2) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on risk of morphometric vertebral fractures and on distal forearm BMD, compared to placebo, in patients who are bisphosphonate-intolerant (i.e. patients with a contraindication or a history of intolerance to bisphosphonates, or those considered by their physician to be unsuitable for bisphosphonate treatment).

(3) **Objective:** To assess the safety and tolerability of treatment with odanacatib 50 mg once weekly, compared to placebo, in patients who are bisphosphonate-intolerant (i.e. patients with a contraindication or a history of intolerance to bisphosphonates, or those considered by their physician to be unsuitable for bisphosphonate treatment).

(4) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on skeletal microarchitecture as assessed by quantitative 2-D histomorphometry and 3-D μ-CT on transilial bone biopsy specimens compared to placebo.

(5) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on the incidence of major atherosclerotic cardiovascular events compared to placebo.

(6) **Objective:** To estimate the treatment effect with odanacatib 50 mg once weekly at 3 years on the incidence of each of the fracture categories (vertebral, hip and non-vertebral).

(7) **Objective:** To determine if baseline protein expression levels (as determined by proteomics) can predict response to treatment with odanacatib 50 mg once weekly on the risk of hip and/or vertebral fracture and/or increase in BMD of the hip and spine.

(8) **Objective:** To determine if change-from-baseline protein expression levels (as determined by proteomics) (< 12 months) can predict long-term response to treatment with odanacatib 50 mg once weekly on the risk of hip and/or vertebral fracture and/or increase in BMD of the hip and spine.

(9) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on the risk of hip, vertebral, and/or non-vertebral fractures compared to placebo as a function of genotype.

(10) **Objective:** To assess the effect of treatment with odanacatib 50 mg once weekly on lumbar spine, total hip, femoral neck, trochanter and distal forearm BMD, height, safety & tolerability, drug exposure, and the levels of circulating protein markers of bone resorption and formation, compared to placebo, as a function of genotype.

(11) **Objective:** To assess the effect of genotype on the risk of hip, vertebral, and/or non-vertebral fractures.

(12) **Objective:** To assess how DNA polymorphisms and protein expression levels relate to appendicular lean body mass assessed by total body DXA.

2.2 SUBJECT/PATIENT INCLUSION CRITERIA

Note that patients inappropriately randomized into the trial (i.e., those who do not meet entry criteria) may be discontinued from the study at the discretion of the SPONSOR.

1) Patient is a woman and is ≥ 65 years of age on the day of Randomization.

2) Patient meets one of the following:

   a) Patient is a candidate for osteoporosis therapy (bisphosphonates, strontium, or PTH), has BMD T-score ≤ -1.5 at either the total hip or femoral neck site, BMD

T-score ≥ -4.0 at both sites, and has one prior vertebral fracture (defined as anterior, mid or posterior height loss of >20%). – OR-

b) Patient is a candidate for osteoporosis therapy (bisphosphonates, strontium, or PTH), has BMD T-score ≤ -2.5 at either the total hip or femoral neck site, BMD T-score ≥ -4.0 at both sites, and does not have a prior vertebral fracture (defined as anterior, mid or posterior height loss of >20%). – OR -

c) Patient is not a suitable candidate for, or has declined osteoporosis therapy (bisphosphonates, strontium, or PTH), has BMD T-score ≤ -1.5 at either the total hip or femoral neck site, and has at least one prior vertebral fracture (defined as anterior, mid or posterior height loss of >20%). – OR-

d) Patient is not a suitable candidate for, or has declined osteoporosis therapy (bisphosphonates, strontium, or PTH), and has a BMD T-score ≤ -2.5 at either the total hip or femoral neck site, and does not have a prior vertebral fracture (defined as anterior, mid or posterior height loss of >20%).

Note: Eligibility for this criterion is based on absolute BMD in g/cm^2 as follows: **Total hip** T-scores of -1.5, -2.5 and -4.0 correspond to 0.759 g/cm^2, 0.637 g/cm^2 and 0.454 g/cm^2 for Hologic machines; 0.820 g/cm^2, 0.694 g/cm^2 and 0.506 g/cm^2 for GE Lunar machines, and 0.736 g/cm^2, 0.614 g/cm^2 and 0.432 g/cm^2 for Norland machines. **Femoral neck** T-scores of -1.5, -2.5 and -4.0 correspond to 0.678 g/cm^2, 0.558 g/cm^2 and 0.378 g/cm^2 for Hologic machines; 0.830 g/cm^2, 0.691 g/cm^2 and 0.483 g/cm^2 for GE Lunar machines, and 0.762 g/cm^2, 0.629 g/cm^2 and 0.430 g/cm^2 for Norland machines.

Note: Patient may not be a suitable candidate for osteoporosis therapy, e.g., due to contraindication, established intolerance, physician's judgment, or patient's unwillingness.

3) The patient has at least one hip that is evaluable by DXA (e.g., contains no hardware from orthopedic procedures).

4) Patient has been postmenopausal for at least 5 years, defined as no menses for at least 5 years OR at least 5 years status post bilateral oophorectomy.

5) Patient understands the study procedures, alternative treatments available and risks involved with the study, and voluntarily agrees to participate by giving written informed consent.

6) Patient is ambulatory.

7) Patient is able to read, understand, and complete questionnaires and diaries.

Note: If a patient who understands the purpose and use of the diary cards and study questionnaires is unable to complete these without assistance, (e.g. due to visual

problems, difficulty writing due to arthritis, inability to read, etc.) a family member or care-giver may assist or may complete the diary card on her behalf.

2.3 SUBJECT/PATIENT EXCLUSION CRITERIA

Note that patients inappropriately randomized into the trial (i.e., those who do not meet entry criteria) may be discontinued from the study at the discretion of the SPONSOR.

1) Patient has chosen treatment with oral bisphosphonates or other agents demonstrated to reduce the risk of hip fracture.

2) Patient has had a prior fragility hip fracture, and she is a suitable candidate for osteoporosis therapy (i.e. bisphosphonates, strontium, or PTH).

   Note: Patient may not be a suitable candidate for osteoporosis therapy, e.g., due to contraindication, established intolerance, physician's judgment, or patient's unwillingness.

3) Patient experienced a clinical fragility fracture (including a clinical vertebral fracture) within 24 months. (Note: Finger, toe, and skull fractures should not be considered with regard to this exclusion criterion.)

   Note: A fragility fracture is defined as a vertebral or non-vertebral fracture, excluding fingers, toes or skull, that occurs when a person falls from a standing height or less, or a fracture sustained without falling such as a vertebral or rib fracture following coughing; these fractures indicate reduced bone strength, as normal-strength bone should be able to withstand this degree of load.

4) Patient has had more than 1 prior vertebral fracture, as defined in Inclusion Criterion 2 above, and she is a suitable candidate for osteoporosis therapy (i.e., bisphosphonates, strontium, or PTH).

5) Patient has or has had evidence of a metabolic bone disorder other than osteoporosis.

6) Patient has a history of renal stones, and serum calcium, serum 25-hydroxyvitamin D and serum PTH are not all within normal limits.

7) Patient has active parathyroid disease. (Note: Serum PTH level should be assessed at screening for patients with a documented history of parathyroid disease. Patients with a history of primary hyperparathyroidism and with curative parathyroidectomy >2 years prior to screening are not excluded.)

8) Patient has a history of thyroid disease not adequately controlled by medication, defined as TSH outside normal limits. Note: If TSH is > 5.5 mU/L and ≤ 8.0 mU/L, patient is eligible if there are currently no plans, (i.e., no clinical need) to change her thyroid medication regimen. In patients with a documented history of thyroid disease, TSH should be assessed at screening.

9) Patient has serum-creatinine >1.6 mg/dL and is considered to have severe renal insufficiency defined as calculated* creatinine clearance ≤ 29 mL/min (National Kidney Foundation K/DOQI Guidelines). *Please note that if serum creatinine ≤ 1.6 mg/dL, there is NO need for calculation of creatinine clearance*.

**Calculated creatinine clearance will be done using the Cockcroft-Gault method for creatinine clearance:* $$Cl_{cr} = \left[\frac{(140 - age)(wt[kg])}{72 \bullet C_s} \right]$$ ** 0.85 with C_s = serum creatinine, and result reported in mL/min.*

Note: Patients with moderate renal insufficiency (creatinine clearance 30-59 mL/min) according to NKF K/DOQI Guidelines may be included provided that serum PTH, serum 25-hydroxyvitamin D, serum calcium and serum phosphorus are *all* within normal limits.

Note: Patients with calculated creatinine clearance ≤ 29 mL/min using the Cockcroft-Gault formula are ineligible *only* in those regions where this is required via documented regulatory request, subsequently approved by the SPONSOR.

10) Patient has received treatment with an agent that has an effect on bone including:

a) estrogen with or without progestin within the prior 6 months (Note: vaginal estrogen creams used not more than 2 times per week are allowed)

b) raloxifene or other SERM (including tamoxifen), tibolone, or an aromatase inhibitor within the prior 6 months

c) sub-cutaneous calcitonin within the prior 6 months (Note: use of intranasal calcitonin either prior to or during the study is permitted)

d) any anabolic steroid use at any time

e) systemic glucocorticoids (≥ 5 mg/day of prednisone or equivalent) for more than 2 weeks in the prior 6 months

f) bisphosphonates: use of any oral bisphosphonate in the 6 months prior to screening; use of any oral bisphosphonate for more than 3 months within the prior 2 years, or lifetime use of more than 6 months total; any lifetime use of IV zoledronate (Note: one dose of IV pamidronate or I.V. ibandronate more than 1 year prior to screening is allowed.)

g) cyclosporin for more than 2 weeks within the prior 6 months

h) fluoride treatment at a dose greater than 1 mg/day for more than 2 weeks at any time

i) strontium-containing products (more than 200 mg elemental strontium daily, including over-the-counter preparations such as OSTEOVALIN™) at any time

j) PTH (1-34 or 1-84) within prior 12 months

k) current use of chemotherapy, or heparin

l) growth hormone at any time

m) cathepsin K inhibitor at any time

n) RANK ligand inhibitor at any time

o) activated Vitamin D (e.g., alphacalcidol) in the prior 3 months

p) current use of vitamin A (excluding beta carotene) >10,000 IU daily, unless willing to discontinue this dose during the study

q) vitamin D supplement > 1200 IU daily, and is unwilling to limit vitamin D supplement to ≤ 1200 IU daily (≤ 8400 IU weekly), with 5600 IU weekly provided as study medication, and up to 400 IU daily permissible as a component of multivitamin

r) protease inhibitors for HIV treatment at any time

s) Patient is taking anti-seizure medication, and indices of calcium metabolism are *not* within normal limits (Note: if indices of calcium metabolism [serum calcium, serum 25-hydroxyvitamin D and serum PTH] are within normal limits, the patient may enroll based on this criterion)

t) current use of systemically administered azole antifungals (for example, ketoconazole, fluconazole, itraconazole, miconazole, posaconozole, ravuconazole, and voriconazole).

11) Patient has a daily calcium intake of <1,200 mg and is unwilling to take study-prescribed supplements, such that her daily calcium intake is approximately 1200 mg.

12) Patient has a history or current evidence of any condition, therapy, lab abnormality or other circumstance that might confound the results of the study, or interfere with the patient's participation for the full duration of the study, such that it is not in the best interest of the patient to participate.

13) Patient has a history of malignancy ≤ 5 years prior to signing informed consent, except for adequately treated basal cell or squamous cell skin cancer or in situ cervical cancer. Patient has had, at any time, a history of melanoma, leukemia, lymphoma, or myeloproliferative disorder.

14) Patient is > 80 years old, AND has a history of recurrent falls (≥ 2 falls in 1 year).

15) Patient is currently participating in or has participated in a study with an investigational compound or device within 30 days prior to signing informed consent.

16) Patient is, at the time of signing informed consent, a user of recreational or illicit drugs or has had a recent history (within the last year) of drug or alcohol abuse or dependence.

17) Patient demonstrates hepatic dysfunction defined as:

Elevation in aspartate aminotransferase (AST) or alanine aminotransferase (ALT) > 3 X Upper Limit of Normal (ULN) *and*

Elevation in total bilirubin > 2 X Upper Limit of Normal (ULN)

18) Patient is considered to be at excessive risk of incident fracture according to local Ethics Committee and/or local regulatory agency.

2.4 STUDY DESIGN AND DURATION

2.4.1 Summary of Study Design

Study Design This is an event-driven, double-blind, randomized, placebo-controlled, multi-center, worldwide study. The patients included in the study will be representative of the general population of postmenopausal women with osteoporosis. The trial will enroll osteoporotic women who have either no prior vertebral fractures or one prior vertebral fracture. All women enrolled in the study will be ≥ 65 years old. Patients without a prior vertebral fracture must have a total hip or femoral neck bone mineral density T-score ≤ -2.5. Patients with a prior vertebral fracture may have a total hip or femoral neck bone mineral density T-score ≤ -1.5. Patients in the 'no prior vertebral fracture' group will be enrolled such that at least 2/3 will be ≥ 70 years old. Patients with more than one pre-existing vertebral fractures, or with a bone mineral density T-score at the total hip or femoral neck <-4.0 will not be eligible unless they are unsuitable for or decline osteoporosis therapy proven effective at non-vertebral sites (i.e., bisphosphonates, strontium, or PTH). Patients may not have had a prior hip fracture, a clinical (symptomatic) vertebral fracture within the past 24 months, and may not have a concomitant illness or laboratory abnormality which might preclude trial completion or confound data interpretation of study results (e.g., history of metabolic bone disease other than osteoporosis). Treatment with bone-active agents (e.g., bisphosphonates, estrogens, etc.) is limited before and during the trial as specified in Section 2.3 and Section 3.2.1.

Odanacatib has not been associated with an increase in skin or respiratory adverse experiences in clinical or preclinical studies. However, because the adverse experiences of morphea-like skin lesions and upper respiratory tract infections have been reported in a phase IIb study of the non-Merck cathepsin K inhibitor balicatib [5], a regulatory agency has requested that the SPONSOR enroll the current trial in two separate phases. This two-phase approach to study enrollment is designed to avoid unnecessarily exposing large numbers of patients to odanacatib, and to permit study of a limited number of patients with close monitoring by an independent Data Monitoring Committee (DMC) to

demonstrate that adverse experiences similar to those seen with balicatib are not associated with odanacatib treatment. In the first phase of enrollment, approximately 1500 patients were randomized in a 1:1 ratio to receive vitamin D_3 (5600 IU once weekly), alone or in combination with odanacatib (50 mg once weekly). Patients also received a sufficient supply of open-label daily calcium supplements, supplied as calcium carbonate, so that their total daily calcium intake (from both dietary and supplemental sources) was approximately 1200 mg. After approximately 1500 patients were enrolled, enrollment was interrupted until all patients in the 'Lead Cohort' received study drug or placebo for at least 9 months. During this 9-month period, data were reviewed by the DMC approximately every 4 months, corresponding to an incremental increase in study drug exposure of approximately 250 patient-years between DMC reviews. After the 9-month safety data were analyzed and reviewed by the DMC as outlined in Section 3.5.5.6, and the risk/benefit ratio was found to be reassuring, the second phase of recruitment began enrolling the balance of ~16,300 patients, the 'Main Cohort'. The primary endpoints in this study are the cumulative incidence of vertebral, non-vertebral and hip fractures which are ordered, as described in Section 3.5.5.1, as follows: First, morphometric vertebral fractures, then second, non-vertebral and hip fractures. The trial is fracture event-driven, and it will end when the pre-specified number of patients with an incident fracture event required for efficacy demonstration has been reached. The numbers of patients with at least one fracture events required to demonstrate fracture risk reduction (assuming efficacy similar to that of alendronate) are 114 at the spine, 237 at the hip, and 824 at non-vertebral sites (Tables 3-13 and 3-14 in Section 3.5.5.4).

Sample Size The precise sample size for the study has not been pre-specified, but is expected to lie between 12,000 and 20,000 study participants depending on the ratio of patients with a prior vertebral fracture to those without a prior vertebral fracture who enroll in the study. According to data obtained from participating investigational sites, the anticipated ratio of patients with a prior vertebral fracture to those without a prior vertebral fracture is approximately 1:3. Based on the overall fracture risk of this mixed population, the sample size is predicted to be approximately 16,300, comprised of approximately 4100 patients with a prior vertebral fracture and approximately 12,200 patients without a prior vertebral fracture. The size of the sample may be revised downward if a higher than anticipated proportion of patients with a prior fracture is enrolled. Conversely, it may be revised upward if a lower than expected proportion of patients with prior fracture is enrolled, or if additional exposure to assess safety is desired.

Trial Duration The study is expected to be approximately 5 years in duration, including 1 year for enrollment, and is not anticipated to exceed a total of 4 years duration for study participants in the 'Main Cohort'. For those in the 'Lead Cohort', trial duration could be approximately 5 years. Interim analyses will be performed after approximately 70% and 85% of hip fracture events have occurred so as to retain the possibility of early trial termination for efficacy (see Data Analysis Section 3.5.5.6). Patients will be monitored for excessive bone loss throughout the study. Those patients determined to have excessive bone loss will be discontinued from the blinded study therapy for treatment with available therapy.

2.4.2 Treatment Plan

2.4.2.1 Blinded Study Therapy and Vitamin D_3

All patients will receive open-label vitamin D_3 (5600 IU once weekly), which will be supplied as two 2800 IU tablets, to be taken once a week. Patients will be randomly assigned to blinded study therapy, either odanacatib (50 mg once weekly) or placebo, and will receive medication according to their assigned treatment group as shown in Table 2-1.

Table 2-1

Sample Allocation Schedule

Treatment Group	Treatment
Odanacatib 50 mg Once Weekly	Odanacatib 50 mg once weekly & 5600 IU Vitamin D_3 once weekly
Placebo	Placebo once weekly & 5600 IU Vitamin D_3 once weekly

Vitamin D_3 and odanacatib will be taken without regard to food. Patients are allowed to choose the day of the week on which to take blinded study therapy and vitamin D_3 supplements (the blinded study therapy and open-label vitamin D_3 need not be taken on the same day of the week). Patients must record on diary cards each day that blinded study therapy and study-provided vitamin D_3 is taken.

Patients should be instructed to take one tablet of blinded study therapy once a week (on their choice of day). As stated above, this may be taken without regard to food. If a dose is forgotten, the patient should take it within 2 days of the scheduled dose and resume taking study therapy on her regular day. Two tablets of blinded study therapy should not be taken on the same day or within a 5 day period as this constitutes an overdose. Compliance with blinded study therapy will be assessed via tablet count and patient report, and should be reinforced at all visits. Sites are encouraged to contact individual patients whom they feel have compliance issues on an as-needed basis.

2.4.2.2 Calcium Supplements

Patients will also receive a sufficient supply of open-label daily calcium supplements, supplied as calcium carbonate, so that their total daily calcium intake (from both dietary and supplemental sources) is approximately 1200 mg but not to exceed 1600 mg. Calcium supplements should preferably be taken with a meal. If a patient wishes to take a calcium supplement other than the one provided by the clinical center, approval must be given by the local clinical monitor.

2.5 LIST OF EFFICACY MEASUREMENTS

The primary outcomes measured in this study are the incidence of first new morphometrically assessed vertebral fracture, first hip fracture and first non-vertebral fracture (for the purposes of this study, non-vertebral fractures exclude fractures of the fingers, toes, face, and skull). Secondary outcomes include increases in bone mineral density at the lumbar spine, the total hip and hip sub-regions, and the distal forearm, as well as changes in biomarkers of skeletal remodeling (resorption and formation).

Spine Fractures: At screening, spine radiographs will be evaluated for the presence or absence of a baseline vertebral fracture using the Genant semi-quantitative scale. In cases in which a vertebral fracture is determined via semi-quantitative methods to have been present at baseline, the spine radiograph will then be evaluated via a quantitative method (i.e., morphometrically). Spine radiographs will be obtained in 100% of patients at baseline, Month 6, and at yearly intervals thereafter. An end-of-study film will also be obtained unless a scheduled film was acquired in the prior 3 months.

It is anticipated that the target number of fracture events will occur when the majority of the patients have undergone 24 to 36 months of treatment. Since enrollment of the 'Lead' and 'Main' Cohorts could take approximately 24 months, the targeted number of fracture events may be achieved approximately 48 to 60 months after study entry for some patients. If this occurs, the 24 month spine radiographs will be used as the primary means for evaluating the incidence of vertebral fractures. For the subset of patients in whom the 24 Month films reveal a vertebral fracture, all prior films will be evaluated using the Genant semi-quantitative technique, as well as via quantitative methods. If the study continues beyond the point at which patients have received at least 36 months of treatment, the Month 36 or the end-of-study (or both) spine radiographs will be evaluated in a similar manner, using Genant semi-quantitative methods first, and then quantitative methods thereafter, for those patients in whom a new vertebral fracture is determined to be present.

Non-Spine Fractures: All other fractures (clinical vertebral, non-vertebral and hip), with the exception of fractures of the fingers, toes, and face, will be assessed clinically and utilizing x-rays for cause throughout the duration of the study. These fractures will be adjudicated to determine whether they are to be classified as fragility (osteoporotic) or non-fragility (trauma-related) fractures. Analyses will be based on adjudicated fractures.

Bone Mineral Density: BMD at the hip and at the lumbar spine will be performed at screening and Randomization respectively, as well as at yearly intervals until the end of the study. BMD assessments will be made in a random, study-wide 10% subset of patients at the total body and distal forearm at Randomization and at yearly intervals until the end of the study. End of study BMD will be assessed unless <6 months have elapsed since the time of the last annual BMD measurement.

Biochemical Markers of Skeletal Turnover: Biochemical markers of bone turnover (s-CTx, u-NTx, s-BSAP and s-P1NP) will be measured in a random, study-wide 10% subset of patients at Randomization, Months 6 and 12, and then yearly until the end of the

study. This will occur in the same subset of patients in whom BMD measurements at all anatomical sites are being performed.

Height: Height will be measured by stadiometer at Randomization as well as at yearly intervals. Stadiometers must be calibrated according to pre-specified procedures. Height measurements may be excluded if obtained with stadiometers that have not been adequately calibrated.

Major Cardiovascular & Cerebrovascular Events: Pre-clinical data suggest that the inhibition of cathepsin K may have a favorable effect on thrombotic cardiovascular and cerebrovascular events due to the stabilization of arterial plaque [6; 7]. There have been recent reports of atrial fibrillation in some patients who received bisphosphonate therapy during clinical trials. Therefore, cardiovascular (CV) events in the categories of thrombotic CV events (including acute and silent myocardial infarction, unstable angina pectoris, and cardiac thrombus), cardiac arrest, cardiac death, and sudden or unexplained death, and new onset atrial fibrillation and atrial flutter, and cerebrovascular events in the categories of ischemic or hemorrhagic strokes and strokes of unknown mechanism will be assessed clinically throughout the duration of the study. These cardiovascular and cerebrovascular events will be adjudicated by an expert panel and analyses will be based solely on adjudicated cardiovascular events.

2.6 LIST OF SAFETY MEASUREMENTS

Clinical evaluations and laboratory measurements including serum chemistry, hematology and urine dipstick will be performed at screening, Months 3, 6, 12, 18, 24, 30, 36, 42, and 48. Adverse experiences will be monitored throughout the study. Excessive bone loss will be monitored by the central BMD QC center and patients will be discontinued from blinded study therapy should this occur (see Section 3.4.1.1). Excessive bone loss is defined as a loss at the lumbar spine or total hip of 7% or greater compared to baseline at any point in the trial. The BMD QC Center will communicate these findings to the SPONSOR and investigator. Bone quality will be assessed by transilial bone biopsies performed at Month 24 and/or Month 36. Compliance with blinded study therapy will be monitored by tablet counts of returned medication.

2.7 DATA ANALYSIS SUMMARY

This section contains a brief summary of the statistical analyses for this study. Full detail is in the Data Analysis Section (DAS) of the protocol details (Section 3.5).

Statistical Methods

Morphometric Vertebral Fractures

Life-table estimates of the percentage of patients with at least one vertebral fracture will be provided for the Month 6 time point and each yearly time point. Treatments will be compared using a generalized linear model for binary data with the complementary log-log transformation of the probability of an event up to the time point. The model will include terms for treatment, stratum (prior/no prior fracture), and geographic region. An

estimate of the hazard ratio from the model will be provided along with its 95% confidence interval. The number (percent) of patients with at least one new morphometric vertebral fracture on the Month 6 and yearly x-rays will also be summarized for each of the treatment groups. The treatment effect at 3 years will also be estimated using life-table estimates. The primary population for the analysis of vertebral fractures will be the Full-Analysis-Set (FAS).

Non-Vertebral, Hip and Clinical Vertebral Fractures

For hip, non-vertebral, and clinical vertebral fractures, Kaplan-Meier estimates of the cumulative incidence of fractures will be graphically displayed. Treatments will be compared using a Cox Proportional Hazard model with terms for treatment, stratum, and geographic region and model-based estimates of the hazard ratio, and its 95% confidence interval, will be provided. The treatment effect at 3 years will also be estimated using Kaplan-Meier estimates. Analyses will be performed using the FAS population.

Stature

The change from baseline will be summarized for each yearly time point. The percentage of patients with a stature loss of more than 1 cm will be summarized and analyzed by a logistic model. The rate of stature loss will be analyzed by a mixed model. Height measurements obtained with stadiometers that have not been adequately calibrated will be excluded from the analysis.

Bone Mineral Density Endpoints

The percent change from baseline in BMD endpoints will be analyzed with a longitudinal model, using the FAS population.

Biochemical Markers

Analysis of log-fraction of baseline value in biochemical markers will utilize the same model as BMD endpoints. Data will be back-transformed for presentation. Biomarkers will be analyzed using the per-protocol population.

Bone Biopsies and Health Resource Utilization and Meal Questionnaires

Bone biopsy, health resource utilization and meal questionnaire data will be summarized for each treatment group.

Major Cardiovascular & Cerebrovascular Events

Time to first major cardiovascular and cerebrovascular event will be summarized in tabular and graphical format by the Kaplan-Meier method.

Population Pharmacokinetics

Population PK analysis and/or PK/PD modeling may be performed. Details will be specified in a separate population PK/PD Statistical Analysis Plan.

Safety Analyses

The primary safety analysis will focus on adverse experiences, with special attention to the Other Non-serious Adverse Experiences mentioned in Sections 3.3.4 and 3.4.1.3. The analysis of adverse experiences will follow a multitiered approach. Safety will be performed using the All-Patients-as-Treated (APaT) population.

Multiplicity

The study will be terminated if the pre-specified number of events is seen for all 3 primary fracture endpoints. The final analysis will be performed at that time. Interim analyses will be performed for the DMC's safety review; the DMC will also monitor efficacy results in the formal efficacy interim analyses as detailed in Section 3.5.5.6 of the Data Analysis Section (DAS).

Multiplicity Due to Multiple Fracture Endpoints

This study is primarily designed to investigate the effect of odanacatib on morphometric vertebral, non-vertebral and hip fractures. To control the false positive error rate due to multiple fracture endpoints in the final analysis, a combination of a step-down closed-testing and Hochberg procedure [8] will be applied. First a step-down procedure will be utilized with the following order of clinical importance: (1) morphometric vertebral fractures, (2) hip and non-vertebral fractures. To control the false positive error rate for multiple tests for hip and non-vertebral fractures, a Hochberg procedure [8] will be used.

In the interim analyses, the trial will not be terminated unless strong evidence (significance) is seen for all 3 primary fracture endpoints, additional detail is provided in the DAS.

Multiplicity Adjustment for Secondary Endpoints

Secondary efficacy endpoints include clinical vertebral fractures, height, BMD measures at different sites, and biochemical markers of bone resorption. Statistical testing for the secondary endpoints will only be performed if the treatment difference for the first primary endpoint (vertebral fractures) is significant. For the purpose of addressing the issue of multiplicity adjustment, BMD measures will be considered as one family. Biochemical markers of bone resorption will be considered another family and clinical vertebral fractures and height a third family. Within each family a Hochberg multiplicity adjustment procedure will be used to ensure a global type I error rate of 5% within each of the groups (conditional on significance at later time points for height and BMD and conditional on significance at earlier time points for biomarkers). No adjustment for multiplicity between the families will be applied. Multiple testing over time will be

handled by a step-down procedure starting from the last time point for BMD and height and from the earliest time point for biomarkers, until the first non-significant time point.

Significance of the analyses of BMD endpoints in the subgroup of (oral) bisphosphonate intolerant patients will only be concluded if there was a significant result for the analysis for the same BMD endpoint at the same timepoint for the full FAS population.

Interim Analyses

A Data Monitoring Committee (DMC) will monitor the safety of the patients in the study on a regular basis and will review efficacy results from the first formal efficacy interim analysis onwards, as detailed in the DAS of the protocol.

Before the first formal efficacy interim analysis, the DMC will mainly review safety and the study should not be terminated early for efficacy. The same efficacy alpha spending function will be used in these interim analyses as for the 2 formal efficacy interims. From the first formal interim analysis onwards, the DMC will review both efficacy and safety and may recommend early termination of the trial for efficacy if strong evidence (significance) is seen for all three primary endpoints. An alpha-spending function approach will be used to handle the multiplicity due to these formal efficacy interim analyses. Details on the actual alpha levels are in the DAS.

Sample Size and Power

Full detail of the sample size and power calculations is in the DAS (Section 3.5).

Sample sizes are based on estimates from the alendronate Fracture Intervention Trial (FIT) trial and are primarily based on the hip fracture data, since the power for non-vertebral and vertebral fractures is higher than for hip fractures. In order to complete recruitment in approximately one year, a flexible sample size depending on the recruited number of patients with and without a prior fracture is used. It is expected that the total sample size will be smaller if more patients with a prior fracture can be recruited. In the calculations it is assumed that the total study duration is approximately 5 years (which is approximately true for the Main Cohort), of which one year is recruitment. Although the actual enrollment distribution and duration are different, these assumptions are considered reasonable as explained in the DAS (Section 3.5). Underlying assumptions of the fracture rates, drop-out rate and relative risks are in Section 3.5. It is expected that the ratio of patients with a prior vertebral fracture to patients without a prior fracture will be approximately 1:3, which means that approximately 4100 patients with a prior fracture and 12,200 patients without a prior fracture will need to be recruited. The size of the sample may be revised downward if a higher than anticipated proportion of patients with a prior fracture is enrolled. Conversely, it may be revised upward if a lower than expected proportion of such patients is enrolled or if additional exposure to assess safety is desired.

The study will be terminated when the pre-specified number of events are seen for all three primary endpoints or when approximately 237 hip fracture events (first fracture), 824 non-vertebral and 114 vertebral fracture events are seen. During the study the

SPONSOR, in consultation with the Steering Committee will monitor the observed fracture events blindly and will determine when sufficient events have accrued for the interim and final analyses.

3. PROTOCOL DETAILS

3.1 RATIONALE

3.1.1 Rationale for This Study

Osteoporosis currently affects approximately 200 million adults worldwide [1]. Approximately 30% of all postmenopausal women in the United States and in Europe have osteoporosis [2]. At least 40% of these women will sustain one or more fragility fractures of the hip, vertebrae, wrist or ribs in their remaining lifetimes. According to data available in 2002, the total annual direct costs for health care attributable to osteoporotic hip fractures was estimated at $18 billion in the US, $5.7 billion in Japan, and $4 billion in the EU [3; 4]. As such, osteoporosis poses a considerable burden both on the health care system and on society at large. The number of patients who will benefit from osteoporosis therapy is growing steadily and is expected to continue to do so for the foreseeable future, since fewer than 25% of osteoporotic patients receive treatment for osteoporosis, the society is aging, and osteoporosis in men is increasingly recognized as a medical issue.

Current treatment options for osteoporosis include bisphosphonates such as alendronate, risedronate and ibandronate, estrogens such as Premarin, selective estrogen receptor modulator (SERMs) such as raloxifene, parathyroid hormone and its analogues, calcitonin, strontium and supplements such as calcium and vitamin D_3. Other than alendronate, zoledronic acid, and hormone replacement therapy, currently available treatments for osteoporosis have demonstrated fracture reduction efficacy only at vertebral and non-vertebral sites. Safety and tolerability limitations exist for most osteoporotic agents. For the bisphosphonates, these include upper GI toxicity and irritation, renal toxicity, and hypothetical concerns about long-term skeletal residence. Osteonecrosis of the jaw in bisphosphonate users has been reported in the published literature. Long term use of estrogens has been shown to increase the risk of myocardial infarction, stroke, fluid retention, deep vein thrombosis and breast cancer. Parathyroid hormone causes concerns because of the observation of osteosarcomas in rats. Consequently there is an unmet medical need for osteoporosis treatments which are highly efficacious but have improved safety and tolerability profiles compared to the currently available agents.

Osteoclastic bone resorption requires demineralization of inorganic bone mineral followed by degradation of organic bone matrix. These processes occur sequentially via two separate mechanisms. The first process involves the secretion of acid into resorption lacunae on the bone surface; the second consists of the degradation of organic (mainly proteinaceous) matrix by cathepsin K. Cathepsin K, which is abundantly expressed in osteoclasts, is a cysteine protease which exhibits collagenolytic activity under acidic conditions. Confirmation of the skeletal activity of cathepsin K in humans comes from the rare hereditary bone disorder, pycnodysostosis, which results from the presence of a defective cathepsin K gene and which is associated with an osteopetrotic phenotype.

Odanacatib is a potent, orally-active inhibitor of cathepsin K which is being developed for the treatment of postmenopausal osteoporosis. Odanacatib has demonstrated robust efficacy in preclinical models, with dose-dependent increases in bone mineral density (BMD) in ovariectomized rabbits and suppression of urinary N-telopeptide cross-links (u-NTx) in ovariectomized (OVX) rhesus monkeys. The combination of increased BMD and suppression of bone resorption markers predicts fracture risk reduction in the clinic. In contrast to currently available anti-resorptive osteoporosis therapies cathepsin K inhibitors are expected to exhibit only limited suppression of bone formation. This hypothesis is supported by observations of normal bone formation in cathepsin K null mice and by preservation of bone formation in ovariectomized rabbits treated with a cathepsin K inhibitor for 6 months. It has also been reported that biochemical markers of bone formation including osteocalcin and bone specific alkaline phosphatase, were not suppressed by the cathepsin K inhibitor balicatib in a 12-month phase IIb study [5]. The absence of suppression of bone formation in tandem with inhibition of bone resorption (a phenomenon referred to as "reduced suppression of bone formation") may provide an advantage over existing anti-resorptive therapies, all of which suppress both resorption and formation.

Odanacatib has several distinguishing characteristics that may confer safety advantages over other cathepsin K inhibitors. Odanacatib is a non-basic molecule and does not accumulate in the acidic cellular compartment of the lysosome; it is also more selective for cathepsin K than for cathepsins B (>200-fold) and L (>1000-fold). This selectivity for cathepsin K is demonstrated in both assays using isolated enzymes and in whole cell assays (which contain lysosomes). By contrast, balicatib retains only 5-6 fold selectivity for cathepsin K over cathepsins B and L in whole cell assays even though it is more selective for cathepsin K in assays using the isolated enzymes. Since cathepsin B has a wide tissue distribution (including the skin) and is involved in apoptosis and collagen turnover, and cathepsin L is mainly involved in epidermal homeostasis, off-target activity of balicatib may explain the cutaneous adverse event profile which has been observed in clinical trials with this drug. Relacatib is non-specific and is equally selective for cathepsins K, L and V. Consequently, off-target activity of relacatib is likely to play a role in the adverse event profile observed. As such, drug-related adverse events seen with balicatib and relacatib are not predicted to arise with odanacatib due to their very different specificity profiles.

Overall, it is anticipated that odanacatib will demonstrate efficacy which is at least similar to that of the bisphosphonates, but without any risk for esophageal irritation. Furthermore, given the rapidity (compared to bisphosphonates) with which odanacatib is cleared from the bone, there is potential for the treatment of younger adult patients (with anorexia nervosa or hypothalamic amenorrhea, for example) whose physicians currently prescribe bisphosphonates with some reluctance.

3.1.1.1 Placebo-Controlled Trials in Osteoporosis

The enrollment of study participants who are at elevated risk for osteoporotic fracture into placebo-controlled osteoporosis trials has been the subject of much recent discussion

[9; 10]. The ethical, scientific and regulatory issues which the SPONSOR took into consideration in designing this trial are reviewed below.

Ethical Rationale

The Declaration of Helsinki and other derivative guidelines uphold the primacy of protecting the rights of research participants, and the Declaration itself makes provision for the use of placebo trials if "compelling and scientifically sound methodologic reasons" can be provided [10]. The SPONSOR considers that there are a number of clear, compelling benefits associated with participation in the proposed trial even for subjects randomized to the placebo arm.

Benefits

"Real-world" Under-treatment of Osteoporosis While standard of care varies considerably around the world, it is common for osteoporosis therapy to be recommended in women demonstrated to have suffered an osteoporotic fracture or who exhibit very low bone mass on bone density testing. While these are the dictates of standard of care, actual "real-world" practice often lags behind. In the U.S., a country in which there is relatively widely available health insurance, only 20% of women already having suffered an osteoporotic hip fracture are prescribed an osteoporosis therapy (including calcium or vitamin D supplements) after their index fracture [11]. Furthermore, women nominally categorized as "treated" for osteoporosis are often not receiving therapy. For example, of those U.S. women actually prescribed osteoporosis therapies such as bisphosphonates, only 20% are still receiving drug 2 years later, and the mean duration of treatment is only 9 months [12]. As a result, up to 80% of women categorized as being "treated" for osteoporosis are receiving no treatment 2 years after therapy was initiated. Under-treatment of osteoporosis is also common in a number of countries where osteoporotic fracture is a reimbursement requirement. Therefore, large numbers of osteoporotic women in the "real world" either are under-treated despite being under medical care or do not receive medical care for osteoporosis at all. This results in many women at high risk for fracture being left without treatment.

By contrast, all women participating in the proposed trial will receive serial bone densitometry, frequent medical monitoring, and calcium plus vitamin D_3 supplements. Half of these study participants will also receive odanacatib which has already been demonstrated to yield increases in bone density comparable to those seen with alendronate.

Vitamin D and Fracture Risk Reduction

In this proposed trial, **all** women will receive supplemental calcium and vitamin D_3 for the duration of the study. Since a number of clinical trials have demonstrated the beneficial effect of calcium plus vitamin D supplementation [13; 14] on fracture risk reduction in osteoporotic women, women in the placebo group of the proposed trial are expected to receive fracture risk reduction benefit as a consequence of study

participation. A meta-analysis of available clinical trial data suggests that a 26% risk reduction in hip fractures may result from administration of 800 IU/d (5600 IU/week) of vitamin D_3 (the proposed dose in this study) [15]. However, it should be appreciated that all previous osteoporosis fracture trials were conducted in vitamin D-replete osteoporosis patients, and that any anti-fracture benefits delivered by the tested drugs were in addition to those provided by nutritional intervention. Whatever the actual fracture risk benefit conferred by vitamin D_3 supplementation in this particular trial, half of study participants will also receive odanacatib and any associated incremental fracture risk reduction for the duration of the trial.

Risks

As of January 2012, odanacatib has been studied in approximately 576 healthy male and female subjects enrolled in 23 phase I studies. In all studies, safety and tolerability were assessed by observation of clinical adverse experiences, laboratory tests, ECG monitoring and physical examinations. Odanacatib has been generally well-tolerated at all doses, including single doses up to 600 mg and multiple oral doses of up to 25 mg daily and 100 mg weekly. There was no indication of causal relationship for any adverse experiences reported and no dose-limiting tolerability issues were identified.

In a Phase IIb dose-ranging study, 399 postmenopausal women were randomized to receive odanacatib at doses of 3, 10, 25 or 50 mg, or placebo weekly. There were 280 study participants who completed 24 months of treatment with odanacatib or placebo. There were no differences between odanacatib and placebo in the incidence of overall clinical adverse experiences, serious adverse experiences, or drug-related serious adverse experiences.

In the first extension of this study (Study Year 3), 189 participants who completed 24 months of treatment were re-randomized to odanacatib 50 mg or placebo once weekly for 12 months. There were 169 patients who completed this extension. Although urinary tract infections (UTIs) AEs were equivalent across treatment groups in Years 1 and 2, an imbalance between the placebo and the 50 mg groups was observed after patients were re-randomized at the end of Year 2. In Year 3, there were a total of 12 patients who experienced UTIs (2 on placebo and 10 on 50 mg odanacatib), with a total of 14 episodes. Urine cultures were only obtained in 3 instances (all in the active treatment group). Results showed only one culture was positive for bacteria while the other two showed no growth. No patient discontinued study therapy due to the AE of UTI and none of the UTI adverse experiences were considered drug related by the investigators.

In the second extension of this study (Study Years 4 and 5) all eligible study participants (N=141) who received odanacatib 3 mg or placebo in the base study (Protocol 004-02) were re-randomized to receive odanacatib 50 mg once weekly during the 24 months of treatment; and participants who received odanacatib 10, 25 or 50 mg in the base study received the same therapy they were assigned to during the 12 month extension (Protocol 004-11) i.e., odanacatib 50 mg once weekly or placebo. As of December 2010, 129 study participants completed Study Year 5. An open-label extension of this study is

ongoing, with a total treatment duration of 10 years being planned. As of December 2011, 115 study participants completed Study Year 6.

At all of its regular meetings, the independent Data Safety Monitoring Committee (DSMC) overseeing the study had recommended that the trial be continued without interruption.

As of September 2011, Protocol 018 has enrolled 16,716 postmenopausal women with osteoporosis, with approximately half of these women having received odanacatib 50 mg once weekly. Study participants have been followed for as long as 42.9 months, with a median follow-up duration of 18.7 months. The independent Data Monitoring Committee (DMC) overseeing this and other odanacatib Phase III studies has met approximately every 4 months since the inception of Protocol 018 in September 2007 and has recommended that the study be continued without interruption. Following its review of available unblinded data in September 2011, the DMC continued to recommend that the trial be continued without interruption. At the Sponsor's request, the DMC also reviewed specific AE categories and recommended that supportive information be collected in all odanacatib trials to confirm the diagnosis of, and to better characterize, specific AEs.

Based on the available safety information and the recommendations of the DMC, the safety profile of odanacatib appears to be favorable thus far.

Women participating in this clinical trial are at elevated risk for incident fracture and, despite the provision of calcium and Vitamin D_3 supplements with or without odanacatib, fracture events will occur in some enrolled patients in both treatment arms. We have, however, excluded women who are at high risk for osteoporotic fracture (e.g. those with BMD T-scores < -4.0, with a prior hip fracture, or with >1 prior vertebral fracture). Consequently, approximately 97% of women randomized to the placebo arm (and a potentially greater proportion of those randomized to odanacatib) would not be expected to experience a hip fracture during the proposed trial; the relatively low risk of incident hip fracture should be properly considered during risk-benefit analysis.

Annual bone density testing at hip and spine anatomical sites will be conducted in 100% of patients in order to identify women whose bone density is decreasing rapidly (e.g. loss at the lumbar spine or total hip of 7% or greater compared to baseline at any point in the trial). Study participants identified as having excessive bone loss will be discontinued from blinded study therapy and will receive open-label 70 mg once weekly alendronate, or 70 mg once weekly alendronate + 2800 IU cholecalciferol, or 70 mg once weekly alendronate + 5600 IU cholecalciferol from the SPONSOR pending primary physician and Ethics Committee assent.

Scientific Rationale

1) Placebo-controlled trials are efficient. The demonstration of fracture risk reduction for odanacatib in a placebo-controlled trial is expected to require approximately

16,300 study participants. By contrast, the demonstration of fracture risk reduction relative to a weak anti-resorptive agent such as raloxifene would require approximately 43,000 women. Such a study is impractical because of the long time (3-4 years) it would require to enroll this large number of patients. A comparator trial of odanacatib versus a more potent anti-resorptive agent such as alendronate would require approximately 125,000 study participants and would take over a decade to recruit (if all prior assumptions about drug efficacy, and event rates in this study population were to be kept constant).

2) Placebo-controlled trials are the best way to ensure the accurate assessment, in absolute terms, of the safety and tolerability of a new chemical entity (NCE). Data derived from comparator trials are considerably more difficult to interpret since adverse events occurring in the setting of NCE administration can only be described in terms of incidence relative to that observed with the comparator. Placebo-controlled safety and tolerability data in a well-designed, well-executed trial are less ambiguous.

Regulatory Rationale

Regulatory agencies have continued to recommend and sometimes require that placebo-controlled fracture-endpoint trials be conducted in support of the registration of new chemical entities for osteoporosis treatment. This practice occurs for the following compelling reasons: 1) Surrogates for bone strength such as bone density and skeletal biomarkers are imperfect predictors of fracture risk. 2) As mentioned above, the accurate, uncomplicated assessment in absolute terms of the safety and tolerability of a new chemical entity is substantially more clear-cut in the context of a placebo-controlled trial; comparator trials are inferior in this regard.

3.1.2 Protection of Study Participants

A variety of measures have been adopted to ensure the safety of participants in this trial.

- Odanacatib has not been associated with skin or respiratory adverse experiences in clinical or preclinical studies. However, because the adverse experiences of morphea-like skin lesions and upper respiratory tract infections have been reported in a phase IIb study of the non-Merck cathepsin K inhibitor balicatib [5], a regulatory agency has requested that the SPONSOR enroll the current trial in two separate phases. This two-phase approach to study enrollment is designed to avoid unnecessarily exposing large numbers of patients to odanacatib, and to permit study of a limited number of patients with close monitoring by a DMC (safety data review approximately every 4 months) to demonstrate that adverse experiences similar to those seen with balicatib are not associated with odanacatib treatment. In the first phase of enrollment, approximately 1500 patients were randomized in a 1:1 ratio to receive vitamin D_3 (5600 IU once weekly), alone or in combination with odanacatib (50 mg once weekly). Patients also received a sufficient supply of open-label daily calcium supplements, supplied as calcium carbonate, so that their total daily calcium intake (from both dietary and supplemental sources) was approximately 1200 mg.

Enrollment was interrupted until all 1500 patients in the 'Lead Cohort' had received study drug or placebo for at least 9 months. On December 2008, the 1500 patients completed 9 months of treatment. During this 9-month period, data were reviewed by the DMC approximately every 4 months, corresponding to an incremental increase in study drug exposure of approximately 250 patient-years between DMC reviews. In January 2009, the 9-month safety data were analyzed and reviewed by an independent DMC, which found the data to be reassuring, and recommended that the balance of ~16,300 patients, the 'Main Cohort', resume enrollment into the trial.

- Treatment with intranasal calcitonin not only prior to but actually *during* the study is permitted. This ensures that use of a therapy broadly approved for the treatment of osteoporosis and demonstrated to reduce fracture risk is possible for all study participants irrespective of the study arm to which each patient has been randomized.

- All study participants will be treated with 5600 IU/week of Vitamin D_3. According to a meta-analysis of available clinical trial data, a 26% risk reduction in hip fractures may result from Vitamin D_3 administration at this dose [15]. Supplemental calcium, demonstrated in placebo-controlled trials to confer fracture protection, will be supplied to participants receiving less than 1200 mg daily.

- Annual bone density testing at hip and spine anatomical sites will be conducted in 100% of patients in order to identify women whose bone density is decreasing rapidly (e.g. loss at the lumbar spine or total hip of 7% or greater compared to baseline at any point in the trial). Study participants identified as having excessive bone loss will be discontinued from blinded study therapy and will receive open-label 70 mg once weekly alendronate, or 70 mg once weekly alendronate + 2800 IU cholecalciferol, or 70 mg once weekly alendronate + 5600 IU cholecalciferol from the SPONSOR pending primary physician and Ethics Committee assent.

- The event-driven study design ensures the shortest trial duration that is permissible by regulatory authorities. No study participant will be exposed to possible placebo for longer than is necessary.

- Study participants with >1 prior vertebral fracture are those who have not tolerated or who have refused to take currently available treatments for osteoporosis. Odanacatib may represent the only therapeutic option for such women.

- DMC oversight of this trial ensures that it will be halted "early" in the event that overwhelming evidence of efficacy is acquired. As such, no patient receiving placebo will be deprived of treatment any longer than is required.

3.1.3 Rationale for Study Population

Odanacatib is being developed to reduce the risk of vertebral, non-vertebral and hip fractures in postmenopausal women who are at elevated risk for osteoporotic fracture. Since fracture risk is well-documented to be a function not only of low bone mass, but also of advanced age and a history of prior fracture [16; 17], the trial will be conducted in

a representative population. The SPONSOR has targeted postmenopausal women who have low bone mass (total hip or femoral neck T-scores ≤ -2.5 for patients without prior vertebral fracture, and total hip or femoral neck T-scores ≤ -1.5 for patients with a prior vertebral fracture), who are over 65 years old. The population will include a significant proportion of patients who have had 1 prior vertebral fracture. This will allow demonstration of efficacy and safety in the population most likely to benefit from this form of treatment.

3.1.4 Summary of Recent Pre-Clinical Data

Odanacatib Research Registration and Imaging Study (ORRIS)

This study was designed to fulfill the non-rodent portion of the preclinical bone quality registration study requirements (in this case, non-human primate). A total of 64 ovariectomized rhesus monkeys were randomized to receive treatment for a duration of 20 months with vehicle, alendronate, or odanacatib at doses 1.7-fold (2mg/kg) and 7-fold (4mg/kg) higher than clinical exposure. Treatment with odanacatib for 18 months resulted in increases in bone mass at typical fracture sites. BMD increases at the lumbar spine of 11.7%, 10.4% and 10.5% relative to vehicle were measured in the odanacatib 2mg/kg, odanacatib 4mg/kg, and alendronate treatment groups, respectively. BMD measurements at the hip showed trends of bone mass increase, although the high variability of these data due to difficulty in positioning the animals with rigid joints seemed to obscure effects resulting in lack of statistical significance. However, volumetric BMD (vBMD) measurements using Quantitative Computed Tomography (QCT), a 3-D measurement technique that is not influenced by positioning, demonstrated an increase in vBMD relative to vehicle in the fracture sites of the femoral neck of 9.7%, 8.9% and 5.6%, and the femoral shaft of 3.1%, 3.1%, 3.9% in the odanacatib 2mg/kg, odanacatib 4mg/kg, and alendronate treatment groups, respectively. DXA measurements of the whole body also showed BMD increases of 8.9%, 7.6%, 6.8%, alluding to a general increase of skeletal bone mass. Daily dosing of odanacatib at 2mg/kg and 4mg/kg reduced bone resorption markers, urinary NTx and serum CTx. Odanacatib unexpectedly displayed an apparent inverse dose-dependence suppression of bone formation markers in non-human primates, i.e. odanacatib 4mg/kg reduced BSAP significantly less than odanacatib 2 mg/kg. Generally, odanacatib reduced the bone formation markers, BSAP and P1NP, to a lesser degree as compared to alendronate. Additionally, distinct from alendronate, odanacatib increased the target engagement marker 1-CTP and the osteoclast specific marker Trap-5b in a dose dependent manner.

3.1.5 Rationale for Dose Regimen

Odanacatib phase I data gathered in healthy postmenopausal women demonstrate that 25 mg and 50 mg of odanacatib administered once weekly result in ~70% and ~60% suppressions of serum C-telopeptides and urine N-telopeptides, respectively, throughout the 7 day dosing interval. Studies were too short in duration (less than or equal to 6 weeks) for meaningful assessment of biomarkers of bone formation. Lower doses of odanacatib resulted in robust suppression of resorption biomarkers (~70%) early in the 7 day dosing interval, but also in sharply attenuated effects on resorption biomarkers

(~20%) in the latter half of the week. No safety signal was detected upon review of unblinded data from these Clinical Pharmacology studies. Based on the reassuring safety profiles associated with all tested doses, and the identical effects of 50 and 100 mg once weekly on biomarkers of bone remodeling in these studies, a top dose of 50 mg was selected for study in phase II trials.

Twenty-four month data from the ongoing phase IIb dose-ranging study conducted in 399 postmenopausal women with low bone mass (T-scores ≤ -2.0) revealed that odanacatib, at the top dose of 50 mg once weekly, resulted in 5.5% and 3.8% increases in lumbar spine and femoral neck BMD from baseline, respectively. By comparison, in historical trials, 24 months of alendronate treatment dosed at 70 mg OW yielded aBMD increases of 5.2% and 2.8%, respectively [18]. Treatment with odanacatib 50 mg once weekly for 36 months yielded continuing BMD increases at the lumbar spine and hip sites that were larger in magnitude than those historically seen with alendronate at these sites in a similar patient population [18].

3.1.6 Rationale for Subject/Patient Genetic Sample Collection

As part of this study, pharmacogenomic analysis may be performed on samples from appropriately consented subjects/patients. The objective of collecting genetic samples in this study is to investigate the relationship between genetic make-up, and the way investigational therapies are absorbed, broken down and eliminated from the body, how they affect the body and how DNA relates to human disease.

The genetic analyses in this study will be conducted to identify the determinants (e.g., single nucleotide polymorphisms, repeats, etc.) of the response of patients to treatment with odanacatib 50 mg once weekly, as assessed by the risk of morphometrically assessed vertebral fractures, hip fractures, and other non-vertebral fractures compared to placebo. Our hypothesis is that treatment with odanacatib reduces the risk of morphometrically assessed vertebral fractures, hip, and other non-vertebral fractures compared to placebo effectively and safely in patients with certain (as-yet-undiscovered) genetic composition. Secondarily, we will seek to determine if there are genetic determinants that associate with baseline, and change-from-baseline, bone mineral density, biochemical markers of bone turnover, 25-hydroxyvitamin D, blood and urine safety/efficacy variables, and other clinical data collected in this study, e.g., fracture.

These genetic markers that predict efficacy and safety in response to treatment with odanacatib will be sought in various ways. The analysis of the nucleic acid data will use genome-wide measurements that seek to identify and confirm loci associated with the aforementioned clinical parameters, e.g., *ctsk*, other cathepsin genes, cyp450 genes, and other genes that affect the absorption, distribution, metabolism, and excretion of odanacatib and similar therapeutics. In addition, single nucleotide polymorphisms (SNPs) at rs4355801, on chromosome 8, near to the TNFRSF11B (osteoprotegerin) gene, and rs3736228, on chromosome 11 in the LRP5 (lipoprotein-receptor-related protein) gene will also be specifically interrogated for possible links to fracture risk and bone mineral density [28]. Also, SNPs in the vitamin D receptor gene will also be specifically interrogated for links to osteoporosis, based on recently published scientific reports [29].

The primary technology used for these measurements will be the Affymetrix Genome-Wide Human SNP Array Version 6.0™ and DMET Plus™ microarrays. The Genome-Wide Array measures, "906,600 SNPs and more than 946,000 probes for the detection of copy number variation." The DMET Plus array assays "markers in all FDA-validated genes, and covers more than 90 percent of the current ADME Core markers as defined by the PharmaADME group." These arrays were developed in collaboration with key scientific leaders from both industry and academia, per the manufacturer's website [30]. In addition, Solexa-type sequencing may be subsequently used to add greater resolution to the array-based data, thereby corroborating and confirming the genotyping findings, if any. Additional descriptions and details on sequencing can be found in [31].

The data generated in this study will be analyzed using a systems biology approach that considers the SNP, the underlying causal genes, and ultimately the association of the target genes to their downstream counterparts. Knowledge of polymorphisms that affect the expression of genes within bone and other tissues will be key to accomplishing this goal. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, gene ontology (GO) terms and genetic networks constructed from F2 mouse crosses will also help to define downstream counterparts for the data set. This analysis does not call for a strict candidate gene approach, because doing so would require all loci of interest to be defined *ad hoc*, and doing so would artificially minimize the available data set.

At the conclusion of this current study, the SPONSOR hopes to have found a set of polymorphisms that can be utilized in subsequent studies to a) enrich for patients more likely to respond to odanacatib 50 mg once weekly and b) identify patients at increased risk for fractures, etc. who would benefit from treatment. If successful, the aforementioned systems biology approach may facilitate an increased understanding of the genetic factors that underlie differences in bone mineral density, fracture, and treatment response to odanacatib. The SPONSOR will ensure that the samples analyzed in this study are put to the best use possible, and will be treated with the utmost respect for the protection of patient privacy. All applicable local and national laws will be followed with regard to data dissemination to patients upon request.

3.1.7 Rationale for Discontinuing Patients Treated with Strong CYP3A4 Inducers

Phase I data from a drug interaction study suggest that concomitant administration of 50 mg odanacatib and 600 mg rifampin (rifampicin) once daily, a strong CYP3A4 inducer, markedly decreases odanacatib plasma concentrations compared to odanacatib given alone. Thus, rifampin could cause a significant reduction in drug exposure levels, thereby compromising the ability to analyze the treatment effect of odanacatib in this trial. In the absence of data with lower doses of rifampin or with other strong CYP3A4 inducers, study participants who initiate chronic treatment with strong CYP3A4 inducers will be discontinued from blinded study therapy. This will be done as a precautionary measure, as it is unknown if chronic use with strong CYP3A4 inducers will cause significant reductions in odanacatib drug exposure levels.

3.2 STUDY PROCEDURES

3.2.1 Concomitant Medication(s)/Treatment

Drugs specifically prohibited in the exclusion criteria are not allowed during the ongoing trial. Listed below are some specific restrictions for concomitant therapy use during the course of the study. If there is a clinical indication for one of these or other medications specifically prohibited during the study, discontinuation from blinded study therapy may be required. The investigator should discuss any questions regarding this with the local clinical monitor. The final decision on any supportive therapy rests with the investigator and/or the patient's primary physician. However, the decision to continue the patient on blinded study therapy requires the mutual agreement of the investigator, the SPONSOR, and the patients.

Bone-Active Agents

Except for the use of intra-nasal calcitonin, patients will not be entered into the study if, during screening, they express an intention to begin estrogen or any other anti-osteoporotic therapy before study completion. Drugs with an effect on bone are prohibited. Soy Isoflavones (e.g., Genistein, Daidzein, Isofem™) are considered estrogens.

If, following Randomization, a patient begins therapy with bisphosphonates, strontium, or PTH, the SPONSOR's Clinical Monitor must be contacted and blinded study therapy will likely be discontinued. In addition, patients who commence therapy with low-dose systemic estrogens for the management of menopausal symptoms *for 6 months or more* must discontinue blinded study therapy. These patients must continue to be followed in the study off-drug. Use of vaginal estrogen up to 2 times per week is permissible during the study.

Anti-fungals

The original version of this protocol included a restriction on the use of systemically or vaginally administered azole antifungals (strong CYP3A4 inhibitors). This was based on results from Phase 1 studies suggesting that ketoconazole, a strong inhibitor of CYP3A4, increases odanacatib pharmacokinetics (PK), with a ~2.4-fold increase in $AUC_{0-\infty}$, while diltiazem (a moderate CYP3A4 inhibitor) modestly increases odanacatib PK, with a ~1.8-fold increase in $AUC_{0-\infty}$. However, the decision to restrict the use of strong CYP3A4 inhibitors such as the azole antifungals was based on narrow preclinical margins established in an odanacatib study of non-human primates conducted at an early stage of the program. Subsequent preclinical studies were conducted where the corresponding safety margins to clinical exposure achieved were ~9-fold in skeletally mature monkeys and ~12-fold in skeletally mature dogs for bone findings, while the preclinical safety margin for soft-tissues was about 12-fold in skeletally immature monkeys. These margins are far higher than the 2.4-fold increase in exposure that could be potentially resulted from concomitant treatment of odanacatib with strong CYP3A4 inhibitors such as ketoconazole.

Several Phase I clinical studies have been conducted to date at odanacatib doses in excess of the 50-mg odanacatib clinical dose, with achieved exposures that were up to ~4.1-fold the clinical exposure. In these studies, odanacatib was generally well tolerated, with adverse events typical of those observed in a Phase I setting, including headache, sore throat, flu-like symptoms, and common cold. Almost all these events were mild or moderate in severity, and all were readily reversible.

In the present study, the initially randomized lead cohort of 1,500 patients (half of whom were randomized to odanacatib) included population PK (pop-PK) sampling. Based on pop-PK modeling, the simulated GMR AUC for patients receiving CYP3A4 inhibitors vs. odanacatib alone is 1.82. This is consistent with the 1.8-fold increase in odanacatib exposure observed with diltiazem co-administration (PN 023), suggesting that the magnitude of CYP3A4-mediated drug-drug interactions are similar in the Phase III and Phase I study populations. In addition, more than 200 Protocol 018 patients to date have taken strong CYP3A4 inhibitors (such as clarythromycin), and more than 1000 patiets have taken moderate inhibitors (such as verapamil, diltiazem and erthytromycin)

An external data monitoring committee (DMC) meets on an ongoing basis every 3-4 months to evaluate the unblinded safety data from the trial, and has consistently recommended that the trial proceed without changes to the protocol. Therefore, based on the aforementioned data, including wider safety margins in preclinical models established after initiation of this study and extensive data collected in healthy volunteers with margins to the clinical dose, we believe that it is no longer necessary to interrupt blinded study therapy for patients using systemically or vaginally administered azole antifungal medications.

Vitamin D

Single-component vitamin D supplements are not permitted. Patients may take multivitamins containing vitamin D as long as the vitamin D dose does not exceed 400 I.U. daily.

Strong CYP3A4 Inducers

If a patient begins a treatment regimen with a strong CYP3A4 inducer (e.g. rifampin, phenobarbital, barbiturates, carbamazepine, phenytoin, St. John's wort, nevirapine, efavirenz, etravirine) during the course of study participation, blinded study therapy should be interrupted for the duration that the strong CYP3A4 inducer is taken, plus an additional 4 weeks to allow CYP3A4 levels to return to baseline. If, however, the duration of treatment with the strong CYP3A4 inducer *exceeds or is expected to exceed 6 months*, the patient must discontinue blinded study therapy. The patient should continue to be followed in the study off-drug. Use of strong CYP3A4 inducers for courses of treatment that last less than 6 months is permitted provided that study drug is interrupted for the duration of treatment with such agents plus an additional 4 weeks as stated above. Intermittent use of CYP3A4 inducers (e.g. intermittent use of barbiturate-

containing medications for headache) is permitted and does not require interruption of study drug.

3.2.2 Diet/Activity/Other

Patients may consume a normal diet, and take their blinded study therapy without regard to food or to other medications.

Calcium supplements should be taken daily, preferably with a meal, and vitamin D_3 supplements taken weekly, without regard to food or other medications.

For patients undergoing transilial biopsies, strenuous physical activity should not be performed during the week following the procedure.

3.2.3 Procedures

3.2.3.1 Informed Consent

3.2.3.1.1 General Informed Consent

The investigator or physician sub-investigator must obtain documented consent from each potential patient in biomedical research or when an investigational drug is administered to the patient in a clinical study, prior to any study related procedures being performed.

Consent must be documented by the patient's dated signature on a consent form along with the dated signature of the person conducting the consent discussion. A copy of the signed and dated consent form should be given to the patient before participating in the trial.

3.2.3.1.2 Consent and Collection of Specimens for Genetic Analysis

During this study, a separate informed consent will be administered for collecting a whole blood specimen for potential future genetic research. Only those subjects/patients who have consented to having this genetic sample collected may have this blood sample drawn. Subjects whose previously collected genetic samples have been inadvertently compromised (e.g., sample integrity, mislabeling, etc) may be invited to re-consent and provide a replacement genetic sample to be included in the genetic analyses.

The investigator or designate is responsible for explaining and verifying the subject's/patient's consent before obtaining such blood samples. It should be explained to the subject/patient that giving the blood sample for genetic information is entirely optional for the subject/patient and participation in the associated clinical study is not dependent upon giving this sample. The approval of the consent form for analysis and the associated protocol procedures (e.g., collection of a blood sample) may, in some cases, proceed independently through Institutional Review Boards, Ethical Review Boards, Independent Ethical Committees, Privacy Committees, etc., from the associated clinical study. In cases where the IRB/ERC approval for the donation of a sample for

genetic analysis is denied or is not accomplished prior to the completion of the clinical study, samples for genetic analysis will not be collected.

Dual-coding of the genetic samples will be performed by the central laboratory and not through the use of genetic identification code (GIC) labels applied by the site. Note that attachment 7 reflects the genetic sampling process used at the beginning of the study until the time of this amendment.

3.2.3.2 Assignment of Baseline Number

A baseline number is assigned to the patient upon signing the informed consent form to identify the patient for all procedures that occur prior to Randomization. A unique baseline number will be assigned to each patient. Baseline numbers must not be re-used for different patients. Any patient who is screened multiple times will retain the original baseline number assigned at the initial screening visit.

3.2.3.3 Stratification

Patients will be stratified to 1 of 2 strata, according to fracture history. Patients without prior vertebral fracture (as assessed centrally) will be assigned to the no prior fracture group. At least 2/3 of the patients in the no prior fracture group will be ≥70 years of age. Patients with a prior vertebral fracture (as assessed centrally) will be assigned to the prior fracture group. The final number of patients in the trial and the numbers in each stratum will depend on the ratio of patients with and without a prior vertebral fracture who enter the study. IVRS (Interactive Voice Recognition System) will be used to keep track of enrollment and the number of patients randomized into each stratum, and will close a stratum as needed.

3.2.3.4 Randomization/Allocation

Each patient will be assigned an allocation number at the time of Randomization. The allocation number will be used to identify the patient for all procedures occurring after Randomization. Once an allocation number is assigned, it can never be re-assigned to another patient.

A single patient/subject cannot be assigned more than 1 allocation number.

3.2.3.5 Monitoring of Enrollment

To be able to have adequate statistical power to demonstrate reduction in hip fracture, approximately 240 hip fractures need to be accrued over the duration of the study. The accrual of this pre-defined number of fracture events is dependent upon the number of study participants enrolled and the overall fracture risk of the study population. The overall fracture risk of the study population is, in turn, dependent on its composition as regards the proportions of patients with and without a prior vertebral fracture. Since the enrollment of patients with or without a prior vertebral fracture will not be restricted, the overall number of patients to be enrolled into the study may be adjusted in response to the composition of the study population, so as to maximize the probability of accruing the requisite number of fracture events.

An illustration of how enrollment monitoring will be operationalized is as follows: A 1,500 patient 'Lead Cohort' will be enrolled into the trial. Recruitment will stop after 1,500 patients are enrolled and study drug or placebo will be administered for 9 months. After the 9-month safety data have been analyzed, have undergone DMC review and have been found to be reassuring, enrollment of the 'Main Cohort' of ~15,000 women will commence. One month after enrollment of the "Main Cohort" has begun, and at monthly intervals thereafter, the number of enrolled patients with and without a prior vertebral fracture will be determined. The overall fracture risk of the population enrolled to that point will be calculated, based on a 3-year hip fracture risk of 1.4% for patients without prior fracture and a 3% hip fracture risk for patients with a prior fracture (see 3.5.5.4 Table 3-13). Based on the proportions of patients with and without a prior vertebral fracture, the required number of patients still to be enrolled can be calculated. The time needed to enroll a study population with an overall fracture risk predictive of the targeted number of hip fractures, and thus the length of the enrollment period, can be also be calculated, based on the enrollment velocity in the preceding months.

3.2.3.6 Treatment

All visit dates subsequent to the Randomization Visit are relative to the date of Randomization.

3.2.3.6.1 Screening (Visit 1)

Prior to coming into the study clinic, potential participants may be pre-screened by imaging (e.g. ultrasound, DXA, x-ray), by phone, chart review, personal interview, physician referral, or available databases, for eligibility criteria that can be assessed in this manner (e.g. age, concomitant therapies, medical conditions, etc.). Those patients who are interested and seem eligible should be invited to the clinic for a formal screening visit. During this visit, study procedures and requirements should be explained and informed consent obtained by the investigator or physician sub-investigator prior to performing any screening procedures. All patients signing a consent form will be given a unique screening number (a "Baseline Number"), generated by the patient database (EDC), that will be used to identify the patient during the screening period for data collection purposes. Each patient will be assigned only one baseline number.

"Candidate for osteoporosis therapy" and Inclusion/Exclusion Criteria

Several entry criteria refer to the prospective study participant's candidacy for osteoporosis therapy (with regard to bisphosphonates, strontium, or PTH). A candidate for osteoporosis therapy is one who has no contraindications to, or is not intolerant of, or would not refuse treatment with bisphosphonates, strontium, or PTH. A patient who is *not* a suitable candidate is one who has contraindications to, or is intolerant of, or would refuse treatment with bisphosphonates, strontium, or PTH (please refer to Operations Manual for further detail).

BMD

Study eligibility regarding BMD criteria will be confirmed by the BMD QA center prior to Randomization. However, sites may determine per local review of the scans that patients with both total hip and femoral neck T-scores > -1.5 are ineligible for the study, and should not send these scans to the QA center or proceed further with screening. Note that lumbar spine BMD will not be considered with regard to study eligibility. Whenever possible, BMD should be performed prior to lateral spine x-rays. Hip BMD scans obtained within 90 days prior to the Randomization visit that comply with the study DXA procedures (as described in the Synarc DXA manual) may be used for study entry.

Note: Spine DXA at screening will *only* be performed in regions where this is required by regulatory agency via documented request, subsequently approved by the SPONSOR.

Lateral Spine Films

Lateral spine films will also be read centrally to determine study eligibility and the stratum to which the patient will be randomized. Lateral spine films obtained within 90 days prior to the Randomization visit that comply with the study x-ray procedures (as described in the Synarc x-ray manual) may be used for study entry.

Safety Laboratory Tests

Lab values that are outside of limits of normal for this population, which the investigator consider to be clinically significant and that are not explained by a clinical diagnosis, should be discussed with the SPONSOR.

Serum Creatinine

Patients with serum creatinine ≤1.6 mg/dL need not undergo further evaluation for renal function, and are eligible for the study.

Patients with serum creatinine >1.6 mg/dL and calculated creatinine clearance ≥60 mL/min are eligible for the study, without further evaluation of renal function.

Patients with severe renal insufficiency (National Kidney Foundation K/DOQI Guidelines) defined as serum creatinine > 1.6 mg/dL and calculated creatinine clearance ≤29 mL/min are not eligible, and further evaluation of renal function should not be performed.

> **Note:** Patients with calculated creatinine clearance ≤29 mL/min using the Cockcroft-Gault formula are ineligible *only* in those regions where this is required via documented regulatory request, subsequently approved by the SPONSOR.

Patients with moderate renal insufficiency, defined as serum creatinine >1.6 mg/dL and calculated creatinine clearance 30 – 59 mL/min must return to the clinic to undergo a blood draw for serum PTH and serum 25-hydroxyvitamin D. These patients may be

included provided that serum PTH, serum 25-hydroxyvitamin D, serum calcium (from screening) and serum phosphorus (from screening) are *all* within normal limits. Calculated creatinine clearance will be done using the Cockcroft-Gault method:

creatinine clearance: $Cl_{cr} = [\ \frac{(140 - age)(wt[kg])}{72 \bullet C_s}\ ]$ * 0.85 with C_s = serum creatinine and result reported in mL/min.

Parathyroid Hormone (PTH) and 25-hydroxyvitamin D

PTH will be measured at screening in patients with a documented history of parathyroid disease.

PTH and 25-hydroxyvitamin D will be measured at screening in patients with a documented history of renal stones, and those with both serum creatinine >1.6 mg/dL and calculated creatinine clearance 30 – 59 mL/min.

PTH and 25-hydroxyvitamin D will be measured at screening in patients taking anti-seizure medication.

Note: PTH and 25-hydroxyvitamin D will also be measured at screening in those regions where this is required via documented regulatory request, subsequently approved by the SPONSOR.

Thyroid Stimulating Hormone (TSH)

TSH will be measured at screening in patients with a documented history of thyroid disease.

Note: TSH will also be measured at screening in those regions where this is required via documented regulatory request, subsequently approved by the SPONSOR.

At the end of the screening visit, all patients who have not been determined to be ineligible will be given vitamin D_3. These patients should also be provided with a sufficient supply of open-label daily calcium supplements, supplied as calcium carbonate, so that their total daily calcium intake from both dietary (assessed by a calcium questionnaire, see Appendix 6.5) and supplemental sources is approximately 1200 mg.

Table 3-1 lists the procedures performed at the screening visit.

Table 3-1

List of Procedures – Screening (Visit 1)

Obtain informed consent for study
Review inclusion/exclusion criteria
Collect medical history[§]
Review prior medications
Perform BMD (total hip and subregions)[‡ ¶ #]
Perform limited physical exam (including vital signs and weight)
Perform lateral spine x-ray[‡]
Collect blood and urine for laboratory safety assessments[†]
Administer calcium questionnaire
Dispense Vitamin D_3 and calcium (for patients with daily calcium intake <1200 mg)
Provide diary card
Additional blood draw for serum PTH and serum 25-hydroxyvitamin D as required.[‖ ††]
[†] Laboratory tests for safety assessments need not be drawn in a fasted state.
[‡] Patients may not be randomized until the central radiologist and BMD QA center confirms study eligibility and determines stratification.
[§] Medical history should include complete information on past/current dental conditions (e.g., periodontal disease, tooth extractions, etc.).
[‖] Additional blood draw for patients with serum creatinine > 1.6 mg/dL and calculated creatinine clearance 30 – 59 mL/min.
[¶] Left hip should be used at screening, unless the left hip is not evaluable. For subsequent hip scans, the same hip should be used as was used at screening.
[#] Spine DXA at screening will *only* be performed in cases where this is required by regulatory agency via documented request, subsequently approved by the SPONSOR.
[††] PTH, 25-hydroxyvitamin D, and TSH will be measured at screening in cases where this is indicated by the protocol, or required by regulatory agency via documented request, subsequently approved by SPONSOR.

3.2.3.6.2 Randomization (Visit 2)

Within 1 month, eligible patients should return to the clinic for a Randomization visit. The remaining Randomization procedures should be completed. Lumbar spine, and total body and distal forearm scans (in the 10% subset) obtained within 3 months prior to the screening visit that comply with the study DXA procedures (as described in the Synarc DXA manual) may be used, and need not be redone at Randomization.

Please note that patients cannot be randomized until the BMD QA center and central radiologist have reviewed DXA and lateral spine films and confirm eligibility and stratification.

PK Sample

PK is obtained in the Lead Cohort only. It is essential to accurately record the date and time that the PK sample is drawn, and enter this into the patient database. Patients will be given the first bottle of blinded study therapy. Patients are allowed to choose the day of the week on which to take the blinded study therapy, and should record each day that blinded study therapy and study-supplied Vitamin D_3 is taken on their diary cards. For patients who have signed an informed consent but are not randomized, the investigator

must submit limited information to the SPONSOR as outlined in the Data Handling Guidelines. Source documentation for nonrandomized patients should remain on file at the site, and is subject to the same retention requirements as data for randomized patients.

Table 3-2 lists the procedures to be performed at the Randomization visit.

Table 3-2

List of Procedures – Randomization (Visit 2)

Review inclusion/exclusion criteria
Perform BMD at the lumbar spine‖
Perform distal forearm and total body BMD§
Perform complete physical exam (including vital signs, height and weight)
Collect blood and urine for laboratory efficacy assessments and archives‡
Collect PK sample in first 1500 patients enrolled
Perform ECG
Dispense study medication (blinded study therapy, vitamin D_3, and calcium for patients with daily calcium intake < 1200 mg)
Provide diary card
Review diary card
Review adverse experiences
Review concomitant medications
Perform Health Resource Utilization (if fracture has occurred)
Collect plasma for proteomics archive
Obtain informed consent for genetic sampling†
Collect genetic sample†
† May be done at Randomization or any subsequent visit.
‡ Efficacy lab tests are performed on a 10% subset of patients. Archives should be collected in all patients. All samples should be fasting specimens, and urine should be second morning void.
§ BMD of the distal forearm and total body performed on a 10% subset of patients.
‖ Patients who had BMD at the lumbar spine performed at screening (Visit 1) do not need to perform this procedure at Randomization (Visit 2).

3.2.3.6.3 Months 3, 9, 18, 30 and 42 (Visits 3, 5, 8, 12, and 16)

PK Sample

PK is obtained in the Lead Cohort only (Months 3 and 9). It is essential to accurately record the date and time that the PK sample is drawn, and enter this into the patient database. In addition, the time (as well as the date) of the two doses of blinded therapy taken prior to the PK draw must be entered.

Meal Questionnaire

It is extremely important to enter into the patient database whether the patient's dose of blinded study therapy prior to the PK draw was taken without food, with a light meal, or with a full meal. As a guideline, "without food" is within 4 hours before, or 30 minutes

after blinded study therapy, a "light meal" is considered a light breakfast (e.g. a piece of toast) or snacks, and a "full meal" is considered a robust/solid breakfast or lunch or dinner. Because PK is only performed in the Lead Cohort, the Meal Questionnaire is only administered in the Lead Cohort as well.

The procedures performed at these visits are summarized in Table 3-3 below.

Table 3-3

List of Procedures – Months 3, 9, 18, 30, and 42[†] (Visits 3, 5, 8, 12, and 16)

Perform limited physical exam (including vital signs and weight)
Collect blood and urine for laboratory safety assessments[‡]
Meal Questionnaire in first 1500 patients enrolled[§]
Collect PK sample in first 1500 patients enrolled[§]
Dispense study medication (blinded study therapy, vitamin D_3, and calcium [if needed])
Provide diary card
Review diary card
Perform tablet count for compliance (only for blinded study therapy)
Review adverse experiences
Review concomitant medication
Perform Health Resource Utilization (if fracture has occurred)
[†] Since this is an event-driven study that will be stopped once a pre-defined number of fracture events have occurred, not all patients may have all visits detailed in the flow-chart.
[‡] Laboratory tests for safety assessments need not be drawn in a fasted state.
[§] Months 3 and 9 only.

3.2.3.6.4 Month 6 (Visit 4)

The procedures performed at this visit are summarized in Table 3-4 below.

Table 3-4

List of Procedures – Month 6 (Visit 4)

Perform limited physical exam (including vital signs and weight)
Perform lateral spine x-ray
Collect blood and urine for laboratory safety assessments[†]
Collect blood and urine for lab efficacy assessments and archives[‡]
Collect PK sample in first 1500 patients enrolled
Meal Questionnaire in first 1500 patients enrolled
Dispense study medication (including blinded study therapy, vitamin D_3, and calcium [if needed])
Provide diary card
Review diary card
Perform tablet count for compliance (only for blinded study therapy)
Review adverse experiences
Review concomitant medication
Perform Health Resource Utilization (if fracture has occurred)
Spine/hip DXA at 6 months will *only* be performed in cases where this is required by regulatory agency via documented request, subsequently approved by the SPONSOR. As with all follow-up DXA measurements, results will be blinded unless EBL criteria are met.
[†] Laboratory tests for safety assessments need not be drawn in a fasted state.
[‡] Efficacy lab tests and archives performed on a 10% subset of patients. These should be fasting specimens, and urine should be second morning void.

3.2.3.6.5 Month 12 (Visit 6)

The procedures performed at this visit are summarized in Table 3-5 below.

Table 3-5

List of Procedures – Month 12 (Visit 6)

Perform limited physical exam (including vital signs, height and weight)
Review concomitant medication
Collect blood and urine for laboratory safety assessments[†]
Collect blood and urine for lab efficacy assessments and archives[‡]
Plasma for proteomics archive
Perform BMD at the total hip (and subregions) and lumbar spine[‖]
Perform BMD at the distal forearm and total body[§]
Perform Health Resource Utilization (if fracture has occurred)
Review adverse experiences
Perform lateral spine x-ray[‡]
Dispense study medication (including blinded study therapy, vitamin D_3, and calcium [if needed])
Provide diary card
Review diary card
Perform tablet count for compliance (only for blinded study therapy)
Administer calcium questionnaire
[†] Laboratory tests for safety assessments need not be drawn in a fasted state.
[‡] Efficacy lab tests are performed on a 10% subset of patients. Archives should be collected in all patients. All samples should be fasting specimens, and urine should be second morning void.
[§] BMD at the distal forearm and total body performed on a 10% subset of patients.
[‖] The same hip must be scanned as was used at baseline.

3.2.3.6.6 Months 15, 21, 27, 33, 39, 45 (Visits 7T, 9T, 11T, 13T, 15T, and 17T)

These are telephone visits, and the procedures conducted are summarized in Table 3-6.

Table 3-6

List of Procedures[†] – Months 15, 21, 27, 33, 39, 45
(Visits 7T, 9T, 11T, 13T, 15T, and 17T)

Review concomitant medications Perform Health Resource Utilization (if fracture has occurred) Review adverse experiences
[†] Since this is an event-driven study that will be stopped once a pre-defined number of fracture events have occurred, not all patients may have all visits detailed in the flow-chart.

3.2.3.6.7 Months 24, 36, 48 (Visits 10, 14, and 18)

The procedures performed at this visit are summarized in Table 3-7 below.

Table 3-7

List of Procedures[†] – Months 24, 36, and 48 (Visits 10, 14, and 18)

Perform BMD at the total hip (and subregions) and lumbar spine[#] Perform BMD at the distal forearm and total body[‖] Perform limited physical exam (including vital signs, height and weight) Perform lateral spine x-ray Collect blood and urine for laboratory safety assessments[‡] Collect blood and urine for lab efficacy assessments and archives[§] Administer calcium questionnaire Dispense study medication (including blinded study therapy, vitamin D_3, and calcium [if needed]) Provide diary card (if needed) Review diary card Perform tablet count for compliance (only for blinded study therapy) Review adverse experiences Review concomitant medication Perform Health Resource Utilization (if fracture has occurred) In patients undergoing transilial bone biopsy at Month 24 and/or Month 36: • Obtain informed consent for transilial bone biopsy approximately 1 month before biopsy • Dispense bone labeling agent approximately 1 month before biopsy • Perform transilial bone biopsy
[†] Since this is an event-driven study that will be stopped once a pre-defined number of fracture events have occurred, not all patients may have all visits detailed in the flow-chart. [‡] Laboratory tests for safety assessments need not be drawn in a fasted state. [§] Efficacy lab tests are performed on a 10% subset of patients. Archives collected in 10% subset except at the end of the study when archives will be collected in everyone. All should be fasting specimens, and urine should be second morning void. [‖] BMD at the distal forearm and total body performed on a 10% subset of patients. [#] The same hip must be scanned as was used at baseline.

3.2.3.6.8 Month 24 (Visit 10) and/or Month 36 (Visit 14): Transilial Bone Biopsy

At participating clinical centers, a transilial bone biopsy will be performed at Month 24 and/or Month 36. Biopsies will be optional for patients and will be done based on a separate informed consent, administered approximately 1 month before the biopsy, prior to any biopsy-related procedures being performed. Patients undergoing this procedure at Month 24 will be provided with an acceptable bone labeling agent at **Month 23** (e.g. demeclocycline or other tetracycline), which will be sourced locally. Patients undergoing this procedure at Month 36 will be provided with an acceptable bone labeling agent at **Month 35**. Patients with bleeding disorders will not be able to undergo a transilial biopsy. Further details of the biopsy procedure, including additional patient requirements for inclusion, will be provided in a separate SOP. It is anticipated that approximately 200 biopsies will be obtained study-wide at Month 24 and approximately 200 biopsies will be obtained at Month 36.

3.2.3.6.9 Visits Past Month 48 (Visit 18)

Since this is an event-driven trial as described earlier in this document, the length of time a patient will be in the trial is unknown, and may extend past the 48 months detailed in the Study Flow Chart. For patients with visits past 48 Months, the following visits and procedures should occur.

- 3 month phone contacts (Months 51, 57, 63, etc.)

The same procedures performed at Month 45 (and similar visits) should be conducted. Please see section 3.2.3.6.6.

- 6 month clinic visits (Months 54, 66, etc.)

The same procedures performed at Month 42 (and similar visits) should be conducted. Please see section 3.2.3.6.3.

- Annual clinic visits (Months 60, 72, etc.)

The same procedures performed at Month 48 (and similar visits) should be conducted. Please see section 3.2.3.6.7.

3.2.3.6.10 End of Study Visit

While the exact timing for the end of the study is unknown and depends on the accumulation of a sufficient number of fracture events, it is estimated to be approximately 3 years after the end of enrollment. The SPONSOR will alert sites when the study will be terminated. At that point, all patients must be brought in for a final visit. Laboratory safety assessment, biomarkers and archive samples do not need to be obtained if less than 30 days have elapsed since the last collection. DXA scans do not need to be performed if less than 6 months have elapsed since the last scan was performed.

Sites will be asked to contact all patients by telephone to schedule the End of Study visits and to re-enforce the importance of this visit. Every effort should be made to ensure that all patients attend the End of Study Visit in the time pre-specified by the SPONSOR.

Table 3-8 lists the procedures to be performed at the final study visit.

Table 3-8

List of End of Study Procedures

Perform BMD at the total hip (and subregions) and lumbar spine[§ ‖ ††]
Perform BMD at the distal forearm and total body[¶]
Complete PE (including vital signs, height and weight)
Perform lateral spine x-ray[§]
Collect blood and urine for laboratory safety assessment[†]
Collect blood and urine for lab efficacy assessment and archives[‡]
Administer calcium questionnaire[#]
Review diary card
Perform tablet count for compliance (only for blinded study therapy)
Review Adverse Experiences
Concomitant Medication Review
Perform Health Resource Utilization (if fracture has occurred)
Plasma for proteomics archive
[†] Laboratory specimen for safety assessments need not be drawn in a fasted state. Do not need to be repeated if less than 30 days elapsed since prior draw.
[‡] Specimens for efficacy test should only be collected in a 10% subset of patients. Archival specimens will be collected for all patients. All samples should be fasting specimens, and urine should be second morning void. Do not need to be repeated if less than 30 days elapsed since prior draw.
[§] BMD to be performed if more than 6 months have elapsed since prior measurement was performed. Lateral spine x-ray to be performed if more than 3 months have elapsed since prior measurement was performed.
[‖] Performed on all patients.
[¶] Performed in 10% of patients.
[#] Administer Calcium Questionnaire if more than 6 months have elapsed since the questionnaire was administered previously.
[††] The same hip must be scanned as was used at baseline.

3.2.3.6.11 Post-study Telephone Call

Approximately 14 days after the last dose of blinded study therapy or study end/discontinuation, the Investigator will contact the patient by telephone to collect serious AE information.

3.2.3.7 Study Termination

Study termination will be the result of the basic scenarios described below:

- *Based upon the advice of the DMC on safety concerns.*

- *Inability to accrue the required number of fracture events*. The Steering Committee and SPONSOR will be monitoring the number of fractures reported during the study, and will advise the SPONSOR if they determine this situation has occurred. A plan for terminating the study will then be identified and communicated to the clinical centers.

- *Based on advice from the DMC if in the formal interim analysis, evidence is shown that odanacatib will be unable to demonstrate efficacy.*

- *Based on advice from the DMC for statistically significant treatment effect seen in one of the formal interim efficacy analyses as detailed in Section 3*.5.

- *Upon accruing the required number of fracture events necessary to test the study's hypotheses*. The Steering Committee and SPONSOR will be monitoring the number of fractures reported in the study, and will determine when the requisite number of fractures have occurred. Every effort should be made to ensure that all patients attend the final study visit in the time pre-specified by the SPONSOR.

3.2.3.8 Blinding/Unblinding

IVRS should be used for emergency unblinding treatment assignment in the event that this is required for patient safety. Every effort should be made to contact the Clinical Monitor prior to such unblinding; however the Investigator may unblind a patient for safety reasons without first contacting the Clinical Monitor. In the event that unblinding has occurred, the circumstances around the unblinding must be documented promptly, and the SPONSOR notified as soon as possible. Note that all patients who have been unblinded must be discontinued from the study.

3.2.3.9 Discontinuation/Withdrawal From Study Therapy

Patients who stop taking study drug prior to the end of the study and agree to continue evaluation and participation in the study must be followed off-drug for the duration of the trial. If a patient discontinues blinded study therapy and/or from study participation, all end of study procedures should be performed.

3.2.3.9.1 Discontinuation Rules

As in all clinical trials, patients will be discontinued from blinded study therapy if the investigator feels that the risk outweighs the benefit of participation for an individual patient. Additionally, the following discontinuation rules will be applied.

- If relationship of the adverse experience to study drug is demonstrated by de-challenge and re-challenge, and the adverse experience is considered severe in intensity, or is serious.

- If a patient has a normal lymphocyte count at baseline, and then during treatment, lymphocyte count falls below the lower limit of the normal range and remains below the normal range on at least 2 repeated measurements, at least 1 week apart, and lymphocyte count returns to the normal range after interruption of study drug.

- If a patient experiences excessive bone loss, defined as a loss at the lumbar spine or total hip of 7% or greater compared to baseline at any point in the trial.

- If a patient demonstrates:

  Persistent elevations [> 3 x Upper Limit of Normal (ULN)] in aspartate aminotransferase (AST) or alanine aminotransferase (ALT) (2 consecutive readings at least 2 weeks apart) - OR -

  Persistent elevations [> 2 x Upper Limit of Normal (ULN)] in total bilirubin (2 consecutive readings at least 2 weeks apart).

- If a patient begins a treatment regimen with a strong CYP3A4 inducer (e.g., rifampin [rifampicin], phenobarbital, barbiturates, carbamazepine, phenytoin, St. John's wort, nevirapine, efavirenz, etravirine) for a period longer than 6 months in duration.

- If a patient commences therapy with low-dose systemic estrogens for 6 months or more.

- If a patient begins treatment with bisphosphonates, strontium, or PTH, the SPONSOR must be contacted and blinded study therapy will likely be discontinued.

Patients will be followed for outcome through the end of the trial even if blinded study therapy is discontinued.

Subjects/patients may withdraw at any time or be dropped from the study at the discretion of the investigator should any untoward effects occur. In addition, a subject/patient may be withdrawn by the investigator or the SPONSOR if he/she violates the study plan or for administrative and/or other safety reasons. The investigator or study coordinator must notify the SPONSOR immediately when a subject/patient has been discontinued/ withdrawn due to an adverse experience (telephone or FAX). When a subject/patient discontinues/withdraws prior to study completion, all applicable activities scheduled for the final study visit should be performed at the time of discontinuation. Any adverse experiences which are present at the time of discontinuation/withdrawal should be followed in accordance with the safety requirements outlined in Section 3.4 SAFETY MEASUREMENTS - DETAILS.

Subjects/patients who donate a blood sample for future genetic analyses may request that their sample be removed from storage and destroyed in accordance with the terms outlined in the informed consent for genetic analyses. Subjects/patients should be informed that withdrawal from the main study does not cause the withdrawal and destruction of the genetic sample. Requests for withdrawal and destruction of the genetic sample should be made in writing to the investigator.

3.2.3.10 Discontinuation/Withdrawal From The Study

Note that patients inappropriately randomized into the trial (i.e. those who do not meet entry criteria) may be discontinued from the study at the discretion of the SPONSOR. Patients who discontinue the study may be permitted to rejoin pending approval of the SPONSOR.

3.3 EFFICACY AND PHARMACOKINETIC MEASUREMENTS

3.3.1 Clinical and Laboratory Measurements for Efficacy

3.3.1.1 Fractures

All reported fractures, other than fractures of the fingers, toes, and face, will be adjudicated centrally via radiology report and/or x-ray. Fractures will be identified as being either osteoporotic, or caused by trauma, stress, or pathology. Analysis of clinical fractures will be based on adjudicated fractures. Adjudication procedures will be detailed in a separate adjudication charter.

At screening, spine radiographs will be evaluated for the presence or absence of a baseline vertebral fracture using the Genant semi-quantitative scale. In cases in which a vertebral fracture is determined via semi-quantitative methods to have been present at baseline, the spine radiograph will then be evaluated via a quantitative method (i.e. morphometrically). Spine radiographs will be obtained in 100% of patients at baseline, 6 Months, and at yearly intervals thereafter. An end-of-study film will also be obtained unless a scheduled film was acquired in the prior 3 months. The central reading of spine films and associated QA procedures will be provided by the vendor in a separate SOP.

It is anticipated that the target number of fracture events will occur when the majority of the patients have undergone 24 to 36 months of treatment. Since enrollment of the 'Lead' and 'Main' Cohorts could take approximately 24 months, the targeted number of fracture events may be achieved approximately 48 to 60 months after study entry for some patients. If this occurs, the 24 month spine radiographs will be used as the primary means for evaluating the incidence of vertebral fractures. For the subset of patients in whom the 24 Month films reveal a vertebral fracture, all prior study films (i.e. those performed at baseline and at Months 6 and 12) will be evaluated using both quantitative and Genant semi-quantitative methodologies. If the study continues beyond the point at which patients have received at least 36 months of treatment, the Month 36 or the end-of-study (or both) spine radiographs will be evaluated in a similar manner, using Genant semi-quantitative methods to identify patients in whom a new vertebral fracture is

determined to be present. In this subset, quantitative assessments of the fractures will then be performed.

All other fractures (clinical vertebral, non-vertebral and hip) will be assessed clinically and utilizing x-rays for cause throughout the study duration.

3.3.1.2 Bone Densitometry

Bone mineral density (BMD) will be measured by dual-energy x-ray absorptiometry (DXA) at the hip (total, femoral neck, trochanter) and lumbar spine for all patients at screening and Randomization, respectively, as well as at yearly intervals until the end of the study. Total body and distal forearm BMD will be measured in a 10% random subset of patients at Randomization and at yearly intervals until the end of the study. End of study BMD will be assessed unless <6 months have elapsed since the time of the last annual BMD measurements. All BMD scans will be centrally evaluated. The BMD QA center will review all baseline scans and confirm patients meet the BMD entry criteria prior to Randomization. The screening Visit 1 hip scan will be used to determine eligibility. Bone densitometry results at the follow-up visits will be blinded to patients and investigators to the extent possible. However, Investigators will be alerted if a patient meets the excessive bone loss criteria (Section 3.4.1.1).

Note: Spine DXA at screening and spine/hip DXA at 6 months will *only* be performed in regions where this is required by regulatory agency via documented request, subsequently approved by the SPONSOR.

All measurements will utilize the same limb and at least 3 vertebrae for all time points (as described in the central vendor DXA manual).

Note: Left hip must be used at screening, unless the left hip is not evaluable. Subsequent hip scans must be performed on the same hip scanned at baseline.

The same machine should be used for a patient throughout the study. Details on DXA acquisition and Quality Assurance will be included in the central vendor DXA manual.

3.3.1.3 Stature

Standing height (without shoes) will be determined by stadiometer at Randomization and annually thereafter. Two measurements will be taken and recorded. If the 2 measurements differ by 4 mm or more, a third and fourth measurement will be obtained. The mean of the last 2 measurements will be used as the estimate of stature. Stadiometers must be calibrated according to pre-specified procedures. Height measurements obtained with stadiometers that have not been adequately calibrated will be excluded from the statistical analysis.

3.3.1.4 Serum and Urine Biochemistry

Biochemical markers of bone metabolism will be measured in a 10% random subset of patients (the same subset of patients receiving DXA at all anatomical sites) at baseline,

Months 6 and 12, and then yearly until the end of the study (see Appendix 6.2). Bone specific alkaline phosphatase (s-BSAP) and serum N-terminal propeptide of Type I collagen (s-P1NP), indices of bone formation, will be measured. N-telopeptides of Type 1 collagen (u-NTx), and C-telopeptides of Type I collagen (s-CTx), indices of bone resorption, will be measured. Indices of calcium and mineral homeostasis (25 hydroxyvitamin D and PTH) will also be measured. These should be collected in the fasting state (no food or drink except water and medications for at least 8 hours), using the second morning void where applicable.

Specific instructions for sampling and processing of these specimens will be provided by the central laboratory.

3.3.1.5 Archival Samples

Serum and urine samples will be obtained and archived from all patients at Randomization, Month 12, and at the end of study. In addition, serum and urine samples will be archived at all time points in the 10% of patients in whom biochemical markers of bone turnover/bone mineral density are measured. These samples should be collected according to the specifications listed in the laboratory manual.

In addition, plasma for the purpose of proteomics archives will be obtained in all patients at Randomization, Month 12, and at the end of study. Proteomics is a systematic large scale analysis of proteins that relies on modern technologies such as Mass Spectrometry, Protein Micro-arrays and Informatics to rapidly identify and quantify proteins. These archival samples may be used to identify novel biomarkers that predict an increased risk of fracture or to explore the patient response to the investigational drug. Methodological approaches will include evaluation of protein expression levels in plasma. These techniques are often referred to as 'open platform' because they measure all of the protein molecules that are present in the plasma sample. Exploratory analysis of these 'expression' patterns can be used to identify sets of molecules that distinguish patient groups such as those with or without a fracture or those who had an optimal or suboptimal response to the investigational therapy. Identification of biomarkers can have great value in helping to target new medicines to those patients who most need them and who are most likely to respond. The best examples of this type of approach to date have been in the field of oncology [19; 20].

3.3.1.6 Pharmacokinetic (PK) Samples

Plasma samples will be collected at Randomization, and Months 3, 6, and 9 for possible population PK analysis. This will be done only in the first 1500 patients ("Lead Cohort") enrolled into the study.

3.3.1.6.1 Meal Questionnaire

At Months 3, 6, and 9, patients will be asked about the intake of their blinded study therapy with respect to food. The dietary assessment should be based on what the patient did on the day she took her last dose of blinded study therapy prior to the PK sample draw. Patients will record on diary cards whether blinded study therapy was taken

without food (within 4 hours before, or 30 minutes after blinded study therapy), with a light meal, or with a full meal. This will be done only in the first 1500 patients ("Lead Cohort") enrolled into the study.

3.3.1.7 Transilial Bone Biopsy

The main objectives of the current study are to assess the effects of odanacatib on fracture risk, BMD, biochemical markers, and clinical safety parameters. Direct assessment of bone quality by histomorphometry and micro CT will be undertaken to gather supportive evidence for the sustained skeletal benefits of long-term odanacatib use.

Unpublished histomorphometric studies in animal models reveal that bone formed during treatment with odanacatib is qualitatively normal. The qualitative assessment of 28 biopsies performed following 24 month treatment in the ongoing Phase II study in postmenopausal women (Protocol 004) did not show any abnormalities, and none of the results on individual specimens departed significantly from the reference database. Giant osteoclasts were not observed. There appeared to be no clinically important differences among treatment groups for activation frequency, bone formation rate or osteoclast surface/bone surface ratio.

In addition to BMD, skeletal microarchitecture is also an important determinant of bone quality. Lamellar bone, the type of bone tissue found overwhelmingly in the adult skeleton, is well-organized, optimally mineralized, and mechanically competent. Woven bone, on the other hand, is deposited when bone is formed rapidly; it is poorly-organized, undermineralized, and mechanically inadequate. In the adult skeleton, woven bone is rare, and is seen principally in fracture callus, Paget's disease of bone, and severe secondary hyperparathyroidism associated with renal osteodystrophy. Treatment of osteoporotic patients with sodium fluoride, an osteotropic agent, also frequently causes woven bone. Assessment of woven and lamellar bone is thus important in evaluating bone quality, both during normal physiological remodeling and during therapeutic intervention.

Available techniques to examine bone structure and assess bone quality include histomorphometry and micro-computed tomography (μCT). Histomorphometry is a destructive 2-dimensional (2D), microanatomic method that assesses structural properties of bone by visualizing individual trabeculae. Structural endpoints are measured on 6 μm sections derived from transilial bone biopsy specimens. Micro-computed tomography (μCT) is a non-destructive, non-in vivo, 3-dimensional (3D), radiation-based imaging method that has sufficient resolution (~20 μm) to visualize individual trabeculae. It can be used ex-vivo to assess 3D structural indices and provide direct assessment of trabecular structure that overcomes the limitations inherent in 2D methodology. 3D μCT can be applied to fresh or plastic-embedded transilial bone biopsy specimens. Though 2D histomorphometry is the more established technique, 3D μCT is likely to provide complementary and possibly more accurate information about bone quality. Both techniques will be used to analyze transilial bone specimens in this study.

Consenting patients at participating clinical centers will undergo a transilial bone biopsy at Month 24 and/or Month 36. While participation in the transilial bone biopsy is optional for patients, all patients should be strongly encouraged to participate in this procedure as this data will provide a direct assessment of bone quality by histomorphometry and micro-computed tomography. These measurements provide gold-standard data regarding skeletal architecture and offer dynamic information such as the bone formation rate.

The details of this procedure and analysis will be addressed in a separate SOP. The first 200 patients who consent to a bone biopsy will have the biopsy performed at Month 24. Subsequent patients will have the biopsy performed at Month 36. Patients who consent to a bone biopsy at Month 24 may also participate in the Month 36 biopsy. It is anticipated that approximately 200 biopsies will be obtained study-wide at Month 24 and approximately 200 biopsies will be obtained at Month 36.

3.3.1.8 Healthcare Resource Utilization

Healthcare resource utilization in patients who experience a fracture will be assessed using the Health Resource Utilization questionnaire. This questionnaire is an eCRF and can be found in the portal for the patient database.

3.3.2 Medication Compliance

Patients should be instructed to take one tablet of blinded study therapy once a week (on their choice of day). This may be taken without regard to food. If a dose is forgotten, the patient should take it within 2 days of the scheduled dose and resume taking study therapy on her regular day. Two tablets of blinded study therapy should not be taken on the same day or within a 5 day period as this constitutes an overdose. Compliance with blinded study therapy will be assessed via tablet count and patient report, and should be reinforced at all visits. Sites are encouraged to contact individual patients whom they feel have compliance issues.

3.3.3 Scientific Advisory Committee, Data Monitoring Committee and Steering Committee

The overall development of the odanacatib program for osteoporosis is conducted in collaboration with a Scientific Advisory Committee (SAC). This is a joint committee composed of Merck and external scientific leaders who convene to provide advice on trial design, statistical analysis, and interpretation of study results for the development of odanacatib for the treatment of osteoporosis.

This study will be conducted under the auspices of a Steering Committee and an independent Data Monitoring Committee (DMC). The composition and scope of these committees is described in Appendix 6.4.

3.3.4 Adjudication Procedures

A number of AE categories will be evaluated by external Central Adjudication Committees (CAC). Since fracture is a primary efficacy endpoint of the trial, all fracture

AEs will be evaluated by an external CAC. Several AE categories will be evaluated by external CACs as part of the assessment of the safety of odanacatib. Publicly presented data on the cathepsin K inhibitor, balicatib, suggest that skin and respiratory AEs are associated with its use in postmenopausal women [5]. As such, both skin and respiratory AEs will be carefully monitored in this trial, and will be evaluated by central adjudication committees. While neither dental AEs nor delayed fracture healing AEs have been reported to be associated with cathepsin K inhibitors, both osteonecrosis of the jaw and delayed fracture healing have been reported with use of some antiresorptive agents. As such, dental and the healing of fractures AEs will be monitored, and will also be evaluated by central adjudication committees. There have been recent reports of atrial fibrillation in some patients who received bisphosphonate therapy during clinical trials. Lastly, pre-clinical data suggest that the inhibition of cathepsin K may exert a favorable effect on thrombotic cardiovascular and cerebrovascular events due to the stabilization of arterial plaque. Thus, as a result of these preclinical data and clinical trial AE reports, thrombotic cardiovascular events, cerebrovascular events and arrhythmic (atrial fibrillation and atrial flutter) events will also be adjudicated in Protocol 018. Specific details regarding endpoint definitions and adjudication procedures will be described in separate charters.

Fractures and Delayed Fracture Union AEs

All clinical fracture events (both vertebral and non-vertebral), with the exception of those of the fingers, toes, and face, will be evaluated by an external Central Adjudication Committee (CAC). As is the case in all fracture endpoint trials, a determination will be made for each incident clinical fracture as to whether it is osteoporotic (defined as fractures that occur in the absence of trauma or in a low impact trauma setting that would not have resulted in fracture in an individual *without* osteoporosis), traumatic (i.e., secondary to excessive force capable of causing a fracture in an individual *without* osteoporosis) or due to another cause (e.g., tumor or stress fracture from repetitive low energy force).

Cases of possible delayed fracture union (fractures in which radiographic evidence of union is not present within 3 months [or within 6 months for tibial and femoral shaft fractures] after the original fracture event) will be reviewed by a CAC.

Fractures and Delayed Fracture Union AEs must be followed until resolution.

Dental AEs

All dental AEs and dental procedures (other than routine cleaning) are to be reported. Suspected cases of "osteonecrosis of the jaw" (ONJ) will be evaluated by an external dental CAC. The investigators should pay special attention to the ONJ diagnosis criteria, especially delayed wound healing longer than 8 weeks. Suspected cases of ONJ must be followed until resolution.

Skin AEs

Skin AEs with skin thickening and hardening suggestive of morphea or systemic sclerosis will be evaluated by an external skin CAC for the presence or absence of morphea-like features.

Respiratory AEs

All respiratory adverse experiences meeting the regulatory definition of serious, except when lung cancer is the only diagnosis, will be evaluated by an external respiratory CAC.

Cardiovascular AEs

Cardiovascular (CV) events in the categories of thrombotic CV events (including acute and silent myocardial infarction, unstable angina pectoris, and cardiac thrombus), cardiac arrest, cardiac death, and sudden or unexplained death, and new onset atrial fibrillation and atrial flutter will be evaluated by an external cardiovascular CAC.

Cerebrovascular AEs

While cathepsin K inhibition is postulated to exert a beneficial effect only on thrombotic cerebrovascular events, it is difficult to discriminate between the various types of cerebrovascular events a priori. As such, cerebrovascular events in the categories of ischemic or hemorrhagic strokes and strokes of unknown mechanism will be evaluated by an external cerebrovascular CAC.

3.4 SAFETY MEASUREMENTS

3.4.1 Clinical and Laboratory Measurements for Safety

Safety will be assessed by a clinical evaluation of AEs and inspection of other study parameters including vital signs, physical examination, and laboratory safety assessment (chemistry, hematology, urinalysis) (see Appendices 6.1 and 6.3 for laboratory tests performed and blood draw volumes). Hematology analysis will include white blood cell and differential count. Sites (patients) may be asked to provide additional details regarding specific AEs based on discussions between the SPONSOR and the DMC, following the DMC's periodic review of study data.

3.4.1.1 Excessive Bone Loss Monitoring

In addition, the BMD QC center will analyze scans on a regular basis, and monitor patients for excessive bone loss. Excessive bone loss will be defined as a loss at the lumbar spine or total hip of 7% or greater compared to baseline at any point in the trial. The QC Center will communicate these findings to the SPONSOR and investigator. A second scan will be obtained whenever possible, and the results of both scans will be averaged to determine whether excessive bone loss has occurred. At the end of the study visit, excessive bone loss will be determined by a single scan, and a confirmatory scan will not be necessary. Patients with significant bone loss should be instructed to

discontinue blinded study therapy and pursue treatment for their osteoporosis from their personal physicians. In countries where allowed per local regulatory guidelines, these patients will be offered a complimentary 1-year supply of 70 mg once weekly alendronate, or 70 mg once weekly alendronate + 2800 IU cholecalciferol, or 70 mg once weekly alendronate + 5600 IU cholecalciferol for optional prescription by their primary care physician. Patients who do discontinue blinded study therapy must continue being followed in the study.

3.4.1.2 Medical History/AEs

A complete medical history will be recorded during the Screening Visit (Visit 1). An interim history and AE inquiry will be made at each office visit thereafter.

3.4.1.3 Other Adverse Experiences

Dental AEs

Localized "osteonecrosis of the jaw", generally associated with tooth extraction and/or local infection, often with delayed healing, has been reported rarely with bisphosphonates. Most reported cases of bisphosphonate-associated osteonecrosis have been in cancer patients treated with high doses of intravenous bisphosphonates. There are no preclinical or clinical data to suggest that cathepsin K inhibition would result in this condition. However, bisphosphonates and cathepsin K inhibitors are potent inhibitors of bone resorption. For this reason, we will closely monitor and report dental adverse experiences (e.g. dental abscesses, tooth loss, periodontal disease, etc.). Cases of suspected osteonecrosis of the jaw will be reviewed by a Central Adjudication Committee (CAC). Suspected cases of ONJ must be followed until resolution.

At screening, information about a patient's history of dental disease, including tooth loss and extractions as well as past or active periodontal disease should be collected. Dental procedures (e.g. tooth extractions, etc.), with the exception of routine cleanings, fillings, and exams, should be recorded, even if these are not associated with an adverse outcome. The reason for the procedure should be recorded. Any diagnosis of osteonecrosis of the jaw should include recent dental history (e.g., recent extraction), the onset date of the dental lesion, and a thorough description of the dental treatments provided. It should also be noted whether osteomyelitis was present. All concomitant medications related to dental AEs and procedures should be recorded.

Patients should be asked about dental AEs at each visit. These should be included with the patient data along with descriptive details on the dental AEs and procedures, and may follow the usual timeframe for reporting study data.

Fracture AEs and Delayed Fracture Union

As this is a fracture trial, we will be closely monitoring the fracture events that occur in this study. All reported fracture events, with the exception of fractures of the fingers, toes, and face, must undergo central adjudication. This will require the submission of radiology reports and/or x-rays to the central radiologist (as needed), as well as a detailed

description of how the fracture occurred. In order to ensure timely adjudication of fracture events, patients should be instructed to call the study center whenever they experience a fracture, and not wait until the next scheduled visit to report the event. These should also be entered into the study database as the site learns about them in order to ensure timely acquisition of associated records and facilitate fracture adjudication.

Fracture AEs must be followed until their resolution. Investigators should evaluate fractures with regard to possible delayed fracture union (fractures in which radiographic evidence of union is not present within 3 months [or within 6 months for tibial and femoral shaft fractures] after the original fracture event). If this is suspected, all associated information should be forwarded with the patient data to be adjudicated.

Respiratory Events

Dose-dependent increases in upper respiratory tract infections have been associated with other, less selective cathepsin K inhibitors [5]. While this has not been observed with odanacatib, respiratory adverse experiences will be closely monitored in this trial. Therefore, all respiratory adverse experiences meeting the regulatory definition of serious, except when lung cancer is the only diagnosis, will be evaluated by an external respiratory CAC.

Cardiovascular Events

Pre-clinical data suggest that inhibition of cathepsin K may have a favorable effect on thrombotic cardiovascular (CV) and cerebrovascular events due to stabilization of arterial plaques. There have been recent reports of atrial fibrillation in some patients who received bisphosphonate therapy during clinical trials. Therefore, cardiovascular events in the categories of thrombotic CV events (including acute and silent myocardial infarction, unstable angina pectoris, and cardiac thrombus), cardiac arrest, cardiac death, and sudden or unexplained death, and new onset atrial fibrillation and atrial flutter will be evaluated by an external cardiovascular CAC

Cerebrovascular Events

Cerebrovascular events in the categories of ischemic or hemorrhagic strokes and strokes of unknown mechanism will be reviewed by an external cerebrovascular CAC.

3.4.1.4 Physical Examination, Height, Weight, and Vital Signs

Physical examinations will be performed in order to ensure the appropriateness of enrolling potential patients and to ensure their safety during the conduct of the study. A complete examination will be performed at Randomization and at the end of the study (final visit). Rectal and urogenital examinations are not required unless indicated by the patient's prior history or clinical symptoms. Limited physical examinations are required at other clinic visits. Abnormal findings noted on the physical examination should be assessed for active medical history conditions and/or AEs.

Standing height will be measured without shoes using a stadiometer. Patients will be weighed without shoes or jackets. Height will be measured at Randomization and annually thereafter.

Weight and vital signs including blood pressure, pulse rate, and respiratory rate will be measured at the appropriate time points outlined in the Study Flow Chart.

3.4.1.5 Blood and Urine Safety Assessments

Analysis of hematology, blood chemistry, and urinalysis specimens will be done by the central laboratory. Collection time points of these specimens are outlined in the Study Flow Chart. Specific instructions for sampling and processing the specimens will be provided in the laboratory manual.

3.4.1.6 ECG

Twelve-lead ECGs will be performed at the Randomization Visit, prior to actually randomizing the patient. Abnormal findings should be assessed for active medical conditions.

3.4.2 Lymphocyte Counts

In response to a request from the FDA, the SPONSOR will report to the agency all absolute lymphocyte counts < 1.0×10^3 microL, even if they are within the normal range and are not considered adverse experiences by the Investigator.

3.4.3 Recording Adverse Experiences

An adverse experience is defined as any unfavorable and unintended change in the structure, function, or chemistry of the body temporally associated with the use of the SPONSOR's product, whether or not considered related to the use of the product. Any worsening (i.e., any clinically significant adverse change in frequency and/or intensity) of a preexisting condition which is temporally associated with the use of the SPONSOR's product, is also an adverse experience.

Changes resulting from normal growth and development which do not vary significantly in frequency or severity from expected levels are not to be considered adverse experiences. Examples of this may include, but are not limited to, teething, typical crying in infants and children, and onset of menses or menopause occurring at a physiologically appropriate time.

Adverse experiences may occur in the course of the use of a Merck product in clinical studies or within the follow-up period specified by the protocol, or prescribed in clinical practice, from overdose (whether accidental or intentional), from abuse, and from withdrawal.

Adverse experiences may also occur in screened subjects/patients during any preallocation baseline period as a result of a protocol-specified intervention including washout or discontinuation of usual therapy, diet, placebo treatment, or a procedure.

Such events will be recorded at each examination on the Adverse Experience Case Report Forms/Worksheets.

3.4.4 Definition of an Overdose for This Protocol

For this protocol, an overdose is defined as:

- ingesting 2 or more tablets of blinded study therapy on the same day or within a 5-day period

If an adverse experience(s) is associated with ("results from") the overdose of test drug, the adverse experience(s) is reported as a serious adverse experience, even if no other criteria for serious are met.

If a dose of test drug meeting the protocol definition of overdose is taken without any associated clinical symptoms or abnormal laboratory results, the overdose is reported as a non-serious Event of Clinical Interest (ECI), using the terminology "accidental or intentional overdose without adverse effect."

All reports of overdose with and without an adverse experience must be reported within 24 hours to one of the individuals listed on the sponsor contact information page found in the Administrative Binder.

3.4.5 Immediate Reporting of Adverse Experiences to the SPONSOR

3.4.5.1 Serious Adverse Experiences

SAEs are to be reported within 14 days following cessation of treatment or discontinuation from the study/end of study, whichever occurs later.

Any serious adverse experience, including death due to any cause, which occurs to any subject/patient entered into this study or within 14 days following cessation of treatment or within the established off therapy follow-up period for safety described in the protocol, whether or not related to the investigational product, must be reported within 24 hours to one of the individual(s) listed on the contact information page.

Additionally, any serious adverse experience considered by an investigator who is a qualified physician to be possibly, probably, or definitely related to the investigational product that is brought to the attention of the investigator at any time outside of the time period specified in the previous paragraph also must be reported immediately to one of the individuals listed on the sponsor contact information page found in the administrative binder.

All subjects/patients with serious adverse experiences must be followed up for outcome.

3.4.5.2 Selected Nonserious Adverse Experiences

These selected non-serious experiences are also known as Events of Clinical Interest (ECI) and must be recorded as such on the Adverse Experience Case Report Forms/Worksheets.

Events of clinical interest for this trial include:

> Persistent elevations [> 3 x Upper Limit of Normal (ULN)] in aspartate aminotransferase (AST) or alanine aminotransferase (ALT) (2 consecutive readings at least 2 weeks apart) or persistent elevations [> 2 x Upper Limit of Normal (ULN)] in total bilirubin (2 consecutive readings at least 2 weeks apart) as determined by way of protocol-specified laboratory testing or unscheduled laboratory testing.*

*Note: These criteria are based upon available regulatory guidance documents. The purpose of the criteria is to specify a threshold of abnormal hepatic tests that may require an additional evaluation for an underlying etiology.

Rash and Skin AEs

During the phase IIb trial for odanacatib, the SPONSOR received information regarding an ongoing assessment of a cutaneous safety signal observed with the use of a non-Merck cathepsin K inhibitor balicatib [5]. The adverse experience has been described as "morphea-like, local skin thickening." For that reason, serious or severe rash and adverse experiences suggestive of morphea will be closely monitored in the odanacatib program. All serious skin AEs, or severe (according to intensity rating) skin AEs, or skin AEs with skin thickening and hardening suggestive of morphea or systemic sclerosis will be promptly reported (within 24 hours) as Events of Clinical Interest. In the reporting of these adverse experiences, all associated information that is considered relevant (e.g. medical history, concomitant therapy, the appearance, distribution, and duration of the rash, any treatments employed, dermatology consultation notes generated, the results of any procedures performed as part of the work-up, and if available, photographs of the lesion, etc.) should be completely documented and submitted with the patient data. Skin cancers are not considered ECIs. In addition, the skin AEs with skin thickening and hardening suggestive of morphea or systemic sclerosis will undergo adjudication.

Skin AEs that do not meet the criteria for ECIs above (i.e. those that are non-serious, and non-severe, or not suggestive of morphea) will follow standard data entry cycle times.

All skin AEs (even those not considered ECIs) that are of clinical importance will be followed up for outcome, especially those that are considered serious, drug related or that lead to discontinuation of study therapy. Follow up will continue through resolution of the AE, even if patients have stopped taking study drug.

3.4.6 Evaluating Adverse Experiences

All adverse experiences will be reported to regulatory agencies, IRB/IECs, and investigators in accordance with all applicable global laws and regulations.

Refer to Table 3-9 for instructions in evaluating adverse experiences.

Table 3-9

An investigator who is a qualified physician, will evaluate all adverse experiences as to:

Maximum Intensity	**Mild**	awareness of sign or symptom, but easily tolerated (for pediatric studies, awareness of symptom, but easily tolerated)
	Moderate	discomfort enough to cause interference with usual activity (for pediatric studies, definitely acting like something is wrong)
	Severe	incapacitating with inability to work or do usual activity (for pediatric studies, extremely distressed or unable to do usual activities)
Seriousness	A serious adverse experience is any adverse experience occurring at any dose that:	
	†**Results in death**; or	
	†**Is life threatening; or** places the subject/patient, in the view of the investigator, at immediate risk of death from the experience as it occurred [Note: This does not include an adverse experience that, had it occurred in a more severe form, might have caused death.]; or	
	†**Results in a persistent or significant disability/incapacity** (substantial disruption of one's ability to conduct normal life functions); or	
	†**Results in or prolongs an existing inpatient hospitalization** (hospitalization is defined as an inpatient admission, regardless of length of stay, even if the hospitalization is a precautionary measure for continued observation. (Note: Hospitalization [including hospitalization for an elective procedure] for a preexisting condition which has not worsened does not constitute a serious adverse experience.); or	
	†**Is a congenital anomaly/birth defect (**in offspring of subject/patient taking the product regardless of time to diagnosis); or	
	Is a cancer; or	
	Is an overdose (Whether accidental or intentional.) Any overdose whether or not associated with an adverse experience must be reported within 24 hours.	
	Other important medical events that may not result in death, not be life threatening, or not require hospitalization may be considered a serious adverse experience when, based upon appropriate medical judgment, the event may jeopardize the subject/patient and may require medical or surgical intervention to prevent one of the outcomes listed previously (designated above by a †).	
Duration	Record the start and stop dates of the adverse experience. If less than 1 day, indicate the appropriate length of time and units	
Action taken	Did the adverse experience cause the test drug to be discontinued?	
Relationship to test drug	Did the test drug cause the adverse experience? The determination of the likelihood that the test drug caused the adverse experience will be provided by an investigator who is a qualified physician. The investigator's signed/dated initials on the source document or worksheet, that supports the causality noted on the AE form, ensures that a medically qualified assessment of causality was done. This initialed document must be retained for the required regulatory time frame. The criteria below are intended as reference guidelines to assist the investigator in assessing the likelihood of a relationship between the test drug and the adverse experience based upon the available information. **The following components are to be used to assess the relationship between the test drug and the AE**; the greater the correlation with the components and their respective elements (in number and/or intensity), the more likely the test drug caused the adverse experience (AE):	
	Exposure	Is there evidence that the subject/patient was actually exposed to the test drug such as: reliable history, acceptable compliance assessment (pill count, diary, etc.), expected pharmacologic effect, or measurement of drug/metabolite in bodily specimen?
	Time Course	Did the AE follow in a reasonable temporal sequence from administration of the test drug? Is the time of onset of the AE compatible with a drug-induced effect?
	Likely Cause	Is the AE not reasonably explained by another etiology such as underlying disease, other drug(s)/vaccine(s), or other host or environmental factors

Relationship to test drug (continued)	**The following components are to be used to assess the relationship between the test drug and the AE: (continued)**	
	Dechallenge	Was the dose of test drug discontinued or reduced? If yes, did the AE resolve or improve? If yes, this is a positive dechallenge. If no, this is a negative dechallenge. (Note: This criterion is not applicable if: (1) the AE resulted in death or permanent disability; (2) the AE resolved/improved despite continuation of the test drug; or (3) the study is a single-dose drug study.)
	Rechallenge	Was the subject/patient reexposed to the test drug in this study? If yes, did the AE recur or worsen? If yes, this is a positive rechallenge. If no, this is a negative rechallenge. (Note: This criterion is not applicable if: (1) the initial AE resulted in death or permanent disability, or (2) the study is a single-dose drug study.) NOTE: IF A RECHALLENGE IS PLANNED FOR AN ADVERSE EVENT WHICH WAS SERIOUS AND WHICH MAY HAVE BEEN CAUSED BY THE TEST DRUG, OR IF REEXPOSURE TO THE TEST DRUG POSES ADDITIONAL POTENTIAL SIGNIFICANT RISK TO THE SUBJECT/PATIENT, THEN THE RECHALLENGE MUST BE APPROVED IN ADVANCE BY THE U.S. CLINICAL MONITOR AND THE INSTITUTIONAL REVIEW BOARD/INDEPENDENT ETHICS COMMITTEE.
	Consistency with Study Drug Profile	Is the clinical/pathological presentation of the AE consistent with previous knowledge regarding the test drug or drug class pharmacology or toxicology?
The assessment of relationship will be reported on the case report forms /worksheets by an investigator who is a qualified physician according to his/her best clinical judgment, including consideration of the above elements.		
Record one of the following:		**Use the following scale of criteria as guidance (not all criteria must be present to be indicative of a drug relationship).**
Yes, there is a reasonable possibility of drug relationship.		There is evidence of exposure to the test drug. The temporal sequence of the AE onset relative to the administration of the test drug is reasonable. The AE is more likely explained by the test drug than by another cause. **Depending on data collection method employed, drug relationship may be further graded as follows:**
	Definitely related	There is evidence of exposure to the test drug. The temporal sequence of the AE onset relative to administration of the test drug is reasonable. The AE is more likely explained by the test drug than by another cause. Dechallenge is positive. Rechallenge (if feasible) is positive. The AE shows a pattern consistent with previous knowledge of the test drug or test drug class.
	Probably related	There is evidence of exposure to the test drug. The temporal sequence of the AE onset relative to administration of the test drug is reasonable. The AE is more likely explained by the test drug than by another cause. Dechallenge (if performed) is positive.
	Possibly related	There is evidence of exposure to the test drug. The temporal sequence of the AE onset relative to administration of the test drug is reasonable. The AE could have been due to another equally likely cause. Dechallenge (if performed) is positive.
No, there is not a reasonable possibility of drug relationship		Subject did not receive the test drug OR temporal sequence of the AE onset relative to administration of the test drug is not reasonable OR there is another obvious cause of the AE. (Also entered for a subject with overdose without an associated AE.) **Depending on data collection method employed, drug relationship may be further graded as follows:**
	Probably not related	There is evidence of exposure to the test drug. There is another more likely cause of the AE. Dechallenge (if performed) is negative or ambiguous. Rechallenge (if performed) is negative or ambiguous.
	Definitely not related	The subject/patient did not receive the test drug. **OR** Temporal sequence of the AE onset relative to administration of the test drug is not reasonable. **OR** There is another obvious cause of the AE.

3.5 DATA ANALYSIS

This section outlines the statistical analysis strategy and procedures. A stand-alone Statistical Analysis Plan (SAP), which details the statistical issues and methods for the study was issued prior to the unblinding of the data for the first interim analysis. The current protocol amendment (018-02) incorporates all details from the stand-alone SAP, according to current Merck guidelines. A separate SAP was created for the PK/PD analyses. The protocol may be further amended if changes are made to the primary or important secondary analyses in the course of the study. Changes to other analyses will be listed in the Clinical Study Report for the study, along with an explanation as to why they occurred and when they occurred.

3.5.1 Responsibility for Analyses/In-House Blinding

The statistical analysis of the data obtained from this study will be the responsibility of the Biostatistics and Research Decision Sciences (BARDS) department of Merck & Co., Inc. This study will be conducted using in-house blinding procedures. For the purpose of the formal efficacy analyses, the official clinical database will not be unblinded until medical/scientific review has been completed, protocol violators have been identified, and data have been declared clean.

A Data Monitoring Committee (DMC) will monitor the safety results in the study on a regular basis, in addition, they will review efficacy results in the formal efficacy interims as detailed in Section 3.5.5.6. Interim analyses will therefore be performed on multiple occasions in this study. The results of interim analyses will only be reviewed by the DMC and will not be shared with Merck & Co., Inc. or any of the investigators. The unblinding of the database at the patient level will be limited to an in-house statistician and statistical programmer (if needed) performing the interim analyses, and limited personnel involved in the process for data base lock and unblinding for the interim analyses. The unblinded personnel will not be involved in any discussions regarding the protocol amendment, if any, other than for safety considerations that may arise in the course of the study, in identification of protocol violators, or data validation efforts once unblinded.

Treatment allocation was based on a computer-generated allocation schedule and was assigned by the Merck Interactive Voice Response System (IVRS). Patients were distributed equally amongst the treatment groups.

3.5.2 Hypotheses

Primary

- Treatment with odanacatib reduces the risk of morphometrically assessed vertebral fractures compared to placebo.
- Treatment with odanacatib reduces the risk of hip fractures compared to placebo.
- Treatment with odanacatib reduces the risk of clinical non-vertebral fractures compared to placebo.

Secondary

- Treatment with odanacatib reduces the risk of clinical vertebral fractures compared to placebo.
- Treatment with odanacatib reduces height loss compared to placebo.
- Treatment with odanacatib increases lumbar spine, total hip, femoral neck, trochanter, and distal forearm BMD compared to placebo.
- Odanacatib is safe and well tolerated compared to placebo.
- Treatment with odanacatib decreases biochemical indices of bone resorption (s-CTx and u-NTx) compared to placebo.
- Treatment with odanacatib increases lumbar spine, total hip, femoral neck, and hip trochanter BMD compared to placebo in bisphosphonate-intolerant patients (defined as patients with a contraindication or history of intolerance to bisphosphonates, or those considered by their physician to be unsuitable for bisphosphonate treatment).

3.5.3 Efficacy/Pharmacokinetics/Safety Variables/Time Points of Interest

Primary Endpoints

The primary focus of this study will be on fracture incidence, specifically morphometic vertebral, non-vertebral and hip fractures. Lateral thoracic and lumbar spine radiographs will be obtained for each patient at baseline, Month 6, Month 12, and yearly thereafter. The first primary endpoint will be the first morphometrically assessed vertebral fracture per patient determined from the 6-month and yearly radiographs, analyzed as interval-censored survival data. All available x-ray data will be included in the analyses and time windows are in appendix 6.7 to define timing.

All reported fractures, other than fractures of the fingers, toes, and face, will be adjudicated centrally via radiology reports and/or x-rays. Clinical fractures will be identified as being either osteoporotic, or caused by trauma, stress, or pathology. Efficacy analyses of clinical fractures will be based on adjudicated osteoporotic fractures.

For clinical hip and non-vertebral fractures, the time to first fracture will be the main metric for analysis. For example, for a patient who has a wrist fracture and later in time a hip fracture, the time to the wrist fracture will be used in the analysis of non-vertebral fractures and the time to the hip fracture will be used in the analysis of hip fractures. The hip fracture of this patient will not be used in the analysis of time to first non-vertebral fractures. Cumulative incidences of fractures for Month 6 and each of the yearly time points will be summarized. Time to first hip and non-vertebral fracture will be the second and third primary endpoints. Analysis of clinical fractures will be based on adjudicated fractures.

Secondary Endpoints

Formal testing will be performed for secondary endpoints with an associated hypothesis. For secondary endpoints without an associated hypothesis, p-values will be provided for informational purposes. The multiplicity adjustment strategy for the formal hypothesis tests is described in Section 3.5.5.3.

Clinical Vertebral Fractures

Time to first (adjudicated osteoporotic) clinical vertebral fracture will be a secondary endpoint and will be handled similarly as hip and non-vertebral fractures.

Height

Two height measurements will be taken and recorded on the eCRF. If the 2 measurements differ by 4 mm or more, a third and fourth measurement will be obtained and recorded on the eCRF. The mean of the last 2 measurements will be used as the estimate of stature.

Change from baseline in stature at each of the yearly timepoints will be secondary endpoints. The number (percentage) of patients with a stature loss of greater than 1 cm during the study will also be analyzed, as will the rate of stature loss.

Height measurements performed on non-calibrated stadiometers or otherwise not obtained according to protocol procedures will be excluded from the analyses. Details on the criteria for exclusion from analysis will be documented in a memo before unblinding the team for the final analysis.

Secondary Bone Mineral Density Endpoints

The percent change from baseline in BMD measurements at the lumbar spine, total hip, femoral neck, trochanter, and 1/3 distal forearm will also be secondary endpoints. BMD measurements will be performed at baseline and each of the yearly timepoints.

For the lumbar spine (posterior/anterior [PA] view) measurements, mean BMD values from at least three evaluable vertebrae (four when available) from L1 to L4 will be used. If a vertebra becomes fractured during the study or was fractured at baseline, BMD data for this vertebra will be excluded from all analyses.

For the hip BMD measurement, the side of the hip measured must be consistent throughout the entire study. If the hip side that is being measured becomes fractured during the study, then data of that hip side will be excluded from the time of the fracture onward.

All analyses of BMD efficacy endpoints will incorporate the longitudinal BMD correction factor as determined by a single quality control center.

BMD at the total body and 1/3 distal forearm will only be performed in a random 10% subset of the randomized patients, identified by the IVRS system

Biochemical Markers of Bone Resorption and Bone Formation

The second morning void urine specimen will be obtained for the 10% subset of patients, at baseline, Month 6, 12 and each of the yearly time points. Determinations will include creatinine and urine N-telopeptides of type 1 collagen. The urinary N-telopeptides/creatinine ratio (NTx/Cr) will be determined. A serum specimen will be obtained at the same time points for measurement of serum C-Telopeptides of Type 1 collagen, serum bone-specific alkaline phosphatase, serum N-terminal propeptides of Type I collagen and parathyroid hormone. 25-hydroxyvitamin D will be measured at baseline, Month 12 and end of the study.

The log-transformed fraction from baseline (calculated by dividing the on-treatment measurement by the baseline measurement and then applying a natural logarithm) at each of the treatment time points will be used to analyze and summarize biochemical markers and indices of calcium and mineral homeostasis. Past experience demonstrated that the log-transformation normalizes the distribution of changes in biochemical markers.

Results will be presented using the original scale after back-transformation. Log-transformed fraction from baseline in biomarkers will be secondary endpoints.

Bone Biopsy Endpoints

Qualitative histomorphometry of transilial bone biopsy specimens will be performed at Month 24 and at Month 36 or the end of the study and will be secondary endpoints. Patient assignment to the time point will be determined in a random manner. Skeletal microarchitecture assessed by 2-D histomorphometry and 3-D μCT transilial bone biopsy measurements at Month 24 and Month 36 (or end of the study) will be an exploratory endpoint.

Exploratory Endpoints

Exploratory BMD endpoints

Total body BMD will be an exploratory endpoint and will be handled the same way as the other BMD endpoints.

Indices of Calcium and Mineral Homeostasis

Indices of calcium and mineral homeostasis will be exploratory endpoints and will be handle the same way as markers of bone formation.

Major Cardiovascular and Cerebrovascular Events

Time to first major cardiovascular or cerebrovascular events is also an exploratory endpoint. A list of the adverse experiences to be considered as major cardiovascular or cerebrovascular events is in Section 3.4.1.3, these events will be adjudicated by a central panel and only those considered by the adjudication committee to pertain to the group of major cardiovascular or cerebrovascular events will be used in the analyses.

Health Resource Utilization Questionnaire

At regular visits during the study, patients who experienced a fracture, will be asked if they received medical care for a fracture and the type of medical care (physician or other healthcare professional office, emergency room visit, hospital admission, nursing home or rehabilitation hospital admission, physical therapy clinic visit).

Meal Questionnaire

At Months 3, 6, and 9, patients will be asked about the intake of their last dose of study medication before the visit with respect to food ('without food', 'with a light meal', 'with a full meal'). This will be done in the first 1500 patients ("Lead Cohort") enrolled into the study only.

PK Analyses:

Population pharmacokinetic samples will be obtained for potential evaluation of drug exposure of MK-00822 in patients receiving oral doses of odanacatib.

Safety and Tolerability

Safety and tolerability will be assessed by a clinical review of all relevant adverse experiences and laboratory safety parameters during the double-blind treatment period.

3.5.4 Analysis Populations

For fracture endpoints a "Full-Analysis-Set" (FAS) population will be used. This will include all randomized patients, who took at least one dose of study medication, with follow-up from Randomization to study termination, without regard to protocol violations, compliance to study drug or early treatment discontinuation. Patients will be counted in the treatment group they were randomized to, irrespective of the treatment they actually received.

All available scans to determine morphometric vertebral fractures, will be used in the analyses. For morphometric fractures, a patient with no scan during the study will not be included in the FAS population. For the final analysis of clinical fractures, a calendar date (cut-off) will be determined when the pre-specified number of fracture events are seen. For the analysis of clinical fractures, patients who did not have a fracture by the time of the cut-off date will be censored at the cut-off date, unless they discontinued the study earlier, then they will be censored at the date of the last contact with the patient (either

telephone contact or study visit). Clinical fractures after the cut-off date will not be included in the analysis.

As a supportive approach, vertebral, non-vertebral and hip fracture analyses will also be performed based on the "Per-Protocol" (PP) population, to investigate the influence of missing data and protocol violators on the treatment effect. This approach excludes patients from the FAS population if they have important protocol deviations. Specific criteria to exclude a patient are identified in appendix 6.7, prior to unblinding of study database and will be based on a patient's compliance to study medication, the use of prohibited previous or concomitant medication, and the presence of secondary diagnoses which may influence the efficacy results. Data will be censored at the date of the last dose of study medication plus 30 days, for patients who did not have a clinical fracture by that time. The per-protocol analysis will not be performed if less than 10% patients (included in the FAS approach) are protocol violators.

A similar FAS population will be used as the analysis approach for BMD endpoints and stature, including all randomized patients who took at least one dose of study medication and have the necessary on-treatment information. Missing data will be handled by the longitudinal model, therefore no imputation of missing data will be necessary. As indicated in appendix 6.7, BMD measurements taken more than 30 days after the last dose of study medication will primarily not be included in the analyses, for results to be consistent with those from phase IIb and future phase III studies. For consistency with the approach for fracture endpoints, a supportive analysis including these measurements taken long after the last dose will also be performed.

The analyses of biochemical markers of bone turnover and formation will be performed using the per-protocol approach. BMD at the total body and distal forearm and biochemical marker endpoints will only be performed for approximately 10% of the randomized patients.

Time to first major cardiovascular and cerebrovascular event will be handled similarly as clinical fractures and will be based on the FAS approach.

Summary tables for health resource and meal questionnaires and bone biopsies will be for the FAS population.

Safety analyses will be based on the "All-Patients-as-Treated" (APaT) population, which includes all randomized patients who received at least one dose of study medication. For safety analyses, patients will be included in the treatment group for the treatment they actually received. For safety analyses, missing data will not be imputed.

3.5.5 Statistical Methods

3.5.5.1 Efficacy Analyses

Morphometric Vertebral Fractures

Life-table estimates of the proportion of patients with at least one fracture will be provided for the Month 6 time point and each of the yearly time points, as well as an estimate of the treatment difference and its 95% confidence interval (CI). In the definition of the life-table estimates a patient who does not have a scan at a timepoint (nor later), defined by the time windows in appendix 6.7, will be considered to be censored immediately after the previous time point, since no additional information on the presence/absence of a fracture is available for this patient after the last available scan.

An interval-censored survival approach will be used to evaluate the treatment effect [21; 22]. A generalized linear model for binary data will therefore be used with the complementary log-log transformation of the probability of an event up to the time point, including all available data from the regularly scheduled x-rays. The model will include terms for treatment, stratum (prior/no prior fracture), and geographic region. An estimate of the hazard ratio from the model will be provided along with its 95% confidence interval. The validity of the proportional hazards assumption will be explored (e.g. using the life-table estimates or by investigation of a similar model with treatment-time point interaction added to the model). Consistency of the treatment effect across strata and regions will also be investigated by summary statistics. These analyses and tables will consider only the first fracture per patient. The number (percent) of patients with at least one new morphometric vertebral fracture on the 6-month and yearly x-rays will also be summarized, for each of the treatment groups.

A sensitivity analysis will be performed in treated patients who received at least one dose of double-blind treatment and did not have any on-treatment lateral spine radiographs, and who are therefore excluded from the primary analysis described above. In this group of patients, the fracture rate will be assumed to be equal to that observed in the placebo group of the patients with spine radiograph (primary analysis). This approach conservatively assumes that there is no treatment effect in the group of patients without on-treatment radiograph. The hazard ratio comparing the two treatment groups (placebo and odanacatib) among patients without a spine radiograph will be combined using meta-analytical methods with the hazard ratio among the patients with radiograph. This will provide an overall estimate of the hazard ratio and 95% confidence interval for all patients (both with and without radiograph) and will allow us to investigate the effect of excluding patients without radiographs from the primary analysis.

Summaries will also be provided in the subgroup of bisphosphonate-intolerant patients (defined as patients with a contraindication or history of intolerance to bisphosphonates, or those considered by their physician to be unsuitable for bisphosphonate treatment).

The proportion of patients with at least one fracture during the first 3 years of the treatment period and the treatment difference, will be estimated using the life-table method.

In addition to these analyses of the first vertebral fracture, a summary will also be provided for the number (percent) of patients with 0, 1, 2, or more than 2 fractures up to Month 6 and each of the yearly time points.

Non-Vertebral, Hip and Clinical Vertebral Fractures

For non-vertebral, hip and clinical vertebral fractures, since the exact date of the fracture is available in the database, the data will be analyzed using continuous time-to-event methodologies. Kaplan-Meier estimates of the cumulative incidence of first fracture at different time points will be provided, as well as an estimate of the treatment difference and its 95% CI, and the data will be graphically summarized with Kaplan-Meier plots. Treatments will be compared using a Cox Proportional Hazard model with terms for treatment, stratum, and geographic region. Model-based estimates of the hazard ratio, and its 95% confidence interval, will be provided. The validity of the proportional hazards assumption will be explored (e.g. by investigation of similar models with treatment-time or treatment-log(time) interaction added). Consistency of the treatment effect across strata and regions will also be investigated by Kaplan-Meier or hazard ratio estimates within stratum/region.

The proportion of patients with at least one fracture during the first 3 years of the treatment period and the treatment difference, will be estimated using the Kaplan-Meier method.

In addition to these analyses of the first fracture, a summary will also be provided for the number (percent) of patients with 0, 1, 2, more than 2 fractures up to Month 6 and each of the yearly time points.

As a sensitivity analysis, similar analyses will be performed for all adjudicated clinical hip fractures, irrespective if they were osteoporotic (defined as fractures that occur in the absence of trauma or in a low impact trauma setting that would not have resulted in fracture in an individual *without* osteoporosis), traumatic (i.e., secondary to excessive force capable of causing a fracture in an individual *without* osteoporosis), stress or pathological. A similar sensitivity analysis will also be performed for all adjudicated non-vertebral, adjudicated vertebral and all adjudicated fractures.

Stature

The mean (and standard error [SE]) baseline and follow-up height, as well as the mean and 95% CI on the change from baseline will be tabulated for each of the treatment groups and yearly time points.

A graphical presentation of the mean (+/-SE) change from baseline in stature over time will also be provided.

In addition, the percentage of patients who experienced a stature loss of greater than 1 cm during the study will be tabulated. A logistic model will be used to compare the two treatment groups, including terms for treatment, geographic region and stratum. The model-based odds ratio, and 95% confidence interval, will be provided as an estimate of the treatment effect.

The rate of stature loss will also be examined, using a mixed model with fixed effects for treatment, geographic region, stratum, treatment-year interaction and random intercept and slope (year). The within patient serial correlation of the values over time will be modeled by an unstructured covariance matrix. The magnitude of the slopes (rate of stature loss) will be summarized from the model, as well as the difference between treatments (mean and 95% CI). More detail on this mixed model is in appendix 6.7.

Bone Mineral Density Endpoints

For each body site and timepoint, BMD will be summarized at baseline and follow-up, by the mean and standard error, together with the mean percent change from baseline and its 95% confidence interval.

The percent change from baseline in BMD endpoints at all during-treatment time points will be analyzed using a longitudinal model with fixed effects for treatment, stratum, geographic region, and treatment-by-time interaction[1]. The within patient serial correlation of the values over time will be modeled by an unstructured covariance matrix. More detail on this longitudinal model is in appendix 6.7.

The treatment effect at each of the on-treatment time points will be assessed by evaluating the within- and between-treatment group Least-Squares means (LS mean) and the associated 95% confidence intervals. Treatment difference will be tested at each of the on-treatment time points and adjustment for multiple testing will be handled as discussed below.

Similar analyses will be performed in the subgroup of bisphosphonate-intolerant patients (defined as patients with a contraindication or history of intolerance to bisphosphonates, or those considered by their physician to be unsuitable for bisphosphonate treatment).

A graphical presentation of the LS mean (+/-SE) change from baseline over time will also be provided.

Biochemical Markers

Analysis of log-transformed fraction of baseline value in biochemical markers of bone resorption (s-CTX, u-NTx) and bone formation (s-BSAP, s-P1NP) will utilize the same longitudinal model as BMD endpoints. The means and the associated 95% CIs will be back-transformed for presentation and the delta method will be used to back-transform

[1] The protocol (018-00) foresaw a random intercept in the model. Due to the unstructured covariance matrix, the random intercept is redundant and therefore omitted from the model.

the between-treatment difference (in weighted LS means) and 95% confidence interval. Formal testing of treatment differences will only be performed for resorption markers. A graphical presentation of the weighted LS mean (+/-SE) change from baseline (after back-transformation) over time will also be provided, as well as a graphical presentation of the mean actual values over time by treatment. Formal testing of treatment differences will only be performed for resorption markers.

Indices of calcium and mineral homeostasis will be summarized in similar fashion to biochemical markers, and no statistical testing will be performed to compare the treatment groups.

In addition, frequency tables will be provided for the percentage of patients with 25-hydroxyvitamin D levels <9, ≥9 to <20, or ≥20 ng/mL at baseline and at Month 12.

Bone Biopsies

The 2-D Histomorphometry and 3-D μ-CT transilial bone biopsy measurements at Month 24 and 36 will be summarized (mean, standard error and 95% CI) for each of the treatment groups. The principle purpose of these analyses is to rule out an increased incidence of qualitative abnormalities (e.g. excessive osteoid, woven bone, mineralization defect, marrow fibrosis), the number (percent) of patients with a qualitative abnormality will therefore be summarized for both treatment groups, along with the Miettinen and Nurminen method [23] based 95% confidence interval on the difference in percentage. The incidence of qualitative abnormalities on the bone biopsies will be considered a safety endpoint rather than an efficacy endpoint.

Meal and Health Resource Utilization Questionnaires

Results of the health resource utilization questionnaire will be summarized at each of the during treatment time points and overall, by means of number (%) of patients who received medical care for a fracture since the last visit and the number (%) patients who received each of the types of medical care (physician or other healthcare professional office, emergency room visit, hospital admission, nursing home or rehabilitation hospital admission, physical therapy clinic visit).

In addition, although originally not foreseen in the protocol, event rates for each type of medical care visit will also be summarized overall and by time point for each treatment group, for all types of visits combined and by type of medical care visit. The event rate will be calculated as the number of patients with at least one medical care visit (per type) divided by the number of patient-years of follow-up.

Furthermore, the number and percent of patients who experienced each type of medical care by fracture type will also be summarized.

Results of the meal questionnaire will be summarized at Months 3, 6, and 9, for the patients of the "Lead Cohort".

Major Cardiovascular and Cerebrovascular Events

Time to first major cardiovascular and cerebrovascular event will be summarized in tabular and graphical format by the Kaplan-Meier method and estimation of the hazard ratio.

Population PK

Population PK analysis and/or PK/PD modeling may be performed. Details will be specified in a separate population PK/PD Statistical Analysis Plan.

Genetic Markers

The purpose of the proposed genetic analyses will be to identify those genetic markers or genotypes that associate with osteoporotic end points within two frameworks: (1) treatment with odanacatib and (2) treatment with placebo. Each consenting patient's genetic material will be genotyped to determine if specific genes are shared by all patients who display a progression of the disease, and/or a reverse based upon intervention. The specific goals are three-fold. First, to identify novel genes that are causal drivers of osteoporosis and can be used as potential targets for future intervention; second, to identify loci that predict disease and disease course; and third, to identify genes/proteins that can be utilized as biomarkers in downstream musculoskeletal studies to reflect a patient's likelihood to respond favorably to odanacatib. Identification of genetic markers that associate with musculoskeletal disease, e.g., osteoporosis, will be achieved by using both the biased approach of drug-metabolizing genes and unbiased genome wide scans. The primary analysis will be cross-sectional using the baseline clinical traits, while the secondary analysis will be longitudinal, using multiple regression models across all patients, in each of the two study arms separately, treating the genotypes as the independent variables and the phenotypes being the dependant variables. New composite measures of musculoskeletal disease, e.g., osteoporosis, are envisioned to be developed that combine multiple measures of bone and muscle health, e.g., BMD, LBM, bone quality, muscle quality, circulating protein markers, and genotype) and utilizes this new composite measure in both the cross-sectional case control studies as well as the longitudinal studies described above. Although a particular composite measure has not been chosen as of yet, one potential method of combining data is shown below. For example, a regression analysis can be conducted by simply treating the three measures (BMD, bone formation markers and bone quality [by biopsy or advanced imaging modality]) as elements of the mixed effect model, similar to the way repeated measures are dealt with, and such model can directly test if genotype is associated with any of these measures singly or through composite crossed terms of these measures. Given the evolving nature of the statistical methodology on GWAS, and biological understanding of musculoskeletal disease, any changes to the statistical methodology and analyses would be outlined and documented in the results memorandum.

Summary of Efficacy Analyses:

A summary of the efficacy analyses performed in this study is in Table 3-10.

Table 3-10

Efficacy Analyses

Efficacy Endpoint	Statistical Method	Analysis Approach	Additional remark
Morphometric Vertebral Fractures			
Time to first fracture	Life-table estimates for each time point	FAS + PP	PP only if >10% PVs.
Time to first fracture	Interval censored analysis (cloglog model)	FAS + PP	PP only if >10% PVs.
Time to first fracture	Summary of number (%) of patients with a fracture up to each time point	FAS + PP	PP only if >10% PVs.
Number of fractures	Summary of number (%) patients with 0, 1, 2, >2 fractures	FAS	
Clinical Fractures (Hip, Non-vertebral and Vertebral)			
Time to first fracture	Kaplan-Meier curve and estimates for each time point	FAS + PP	PP only if >10% PVs.
Time to first fracture	Cox proportional hazards model	FAS + PP	PP only if >10% PVs.
Number of fractures	Summary of number (%) patients with 0, 1, 2, >2 fractures	FAS	
Stature			
Change from baseline at each time point	Summary table and line plot over time.	FAS	
Stature loss of >1 cm	Summary of number (%) patients and logistic model	FAS	
Rate of stature loss	Mixed model with fixed effects for treatment, stratum, region, treatment-time and random intercept and slope.	FAS	
Bone Mineral Density Endpoints			
Percent change from baseline at all time points.	Summary table.	FAS	
Percent change from baseline at all time points.	Longitudinal approach and line plot of LS means over time	FAS	
Bone Biopsy Endpoints			
Bone biopsy measurements	Summary table.	FAS	

Efficacy Analyses (Cont.)

Efficacy Endpoint	Statistical Method	Analysis Approach	Additional remark
Qualitative abnormality[†]	Summary of number (%) of patients	FAS	
Biochemical Markers of Bone Resorption and Bone Formation			
Log-transformed fraction from baseline	Summary table of geometric mean percent change from baseline.	PP	
Log-transformed fraction from baseline	Longitudinal model and line plot of LS mean over time.	PP	Testing only for resorption markers.
Indices of Calcium and Mineral Homeostasis			
Log-transformed fraction from baseline	Summary table of geometric mean percent change from baseline.	PP	
Major Cardiovascular and Cerebrovascular Events			
Time to first event	Kaplan-Meier curve and estimates for each time point	FAS	
Health Resource Utilization Questionnaire			
Patients who received medical care and type of care	Summary of number (%.) of patients overall and per time point. Summary of event-rate per type of medical care visit. Summary of number (%) of patients by type of fracture.	FAS	
Meal Questionnaire			
Meal questionnaire items	Summary of number (%) of patients per time point	FAS	Only performed for "Lead Cohort".

FAS= Full Analysis Set; PP= Per Protocol; PV= Protocol Violator
[†] Incidence of qualitative abnormalities on bone biopsies will be considered a safety endpoint.

3.5.5.2 Safety Analyses

Safety and tolerability of odanacatib compared to placebo will be assessed by a clinical and statistical review of all safety data collected throughout the study. The primary safety analysis will focus on adverse experiences, with special attention to the Other Non-serious Adverse Experiences mentioned in Sections 3.3.4 and 3.4.1.3. More detail on the safety analyses is in appendix 6.7.

The primary analysis of adverse experiences will include all events from the start of prime therapy until the end of the trial (cut-off date or last contact date, for patients who discontinued early from the trial before the cut-of date), irrespective of compliance to treatment. A sensitivity analysis, including all adverse experiences that occurred after the start of double-blind treatment and within 3 months (91 days) after any intake of double-blind study medication will also be performed. This sensitivity approach will be used for the Tier 1 tables and selected Tier 2 tables.

The analysis of adverse experiences will follow a multi-tiered approach (Table 3-11). The tiers differ with respect to the analyses that will be performed. Safety parameters or adverse experiences of special interest that are identified *a priori* constitute Tier 1 safety endpoints that will be subject to inferential testing for statistical significance with p-values and 95% CIs provided for between-group comparisons. Other safety parameters will be considered Tier 2 or Tier 3. Tier 2 parameters will be assessed via point estimates with 95% CIs provided for between-group comparisons; only point estimates by treatment group are provided for Tier 3 safety parameters.

AEs (specific terms as well as system organ class terms) and predefined limits of change in laboratory will be classified as belonging to "Tier 2" or "Tier 3," based on the number of events observed.

Tier 1 events include: any skin adverse experiences reported as ECI, any serious respiratory infections confirmed by adjudication, osteonecrosis of the jaw confirmed by adjudication, skin changes related to morphea or scleroderma confirmed by adjudication, and delayed fracture union confirmed by adjudication, provided their incidence is ≥4 patients in one of the treatment groups; otherwise these events will be considered Tier 3.

Tier 2 events include: Any clinical AE, serious clinical AEs, drug-related clinical AEs, and AEs leading to discontinuation of study medication, if their incidence is at least 1% in either of the treatment groups.

The same categories (any events, related events, serious events, discontinuations and events with incidence ≥1%) will also be considered Tier 2 for laboratory events

In addition, specific dental disorders, specific skin disorders and specific respiratory disorders with incidence ≥1% will be handled as Tier 2. Furthermore, hypocalcemia (including decreased blood calcium) will also be considered Tier 2 events.

A predefined subset of the Merck-MedDRA dictionary determined before unblinding the data for the DMC interim analyses will be used to identify skin, dental, and respiratory adverse. This list will be updated regularly by the clinical team in a blinded fashion to account for MedDRA dictionary updates.

Specific dental disorders, specific skin disorders, specific respiratory disorders, and specific AEs with >0% incidence (irrespective of the incidence) will be tabulated as Tier 3 events.

The threshold of at least 4 events was chosen because the 95% CI for the between-group difference in percent incidence will always include zero when treatment groups of equal size each have less than 4 events and thus would add little to the interpretation of potentially meaningful differences. Because many 95% CIs may be provided without adjustment for multiplicity, the confidence intervals should be regarded as a helpful descriptive measure to be used in review, not a formal method for assessing the statistical significance of the between-group differences in AEs and predefined limits of change. For Tier 1 events, estimates of the difference in proportions of patients will be provided, along with the 95% confidence intervals and a formal test for the between-group differences using Miettinen and Nurminen method, taking the differential observation period per patient into account [23]. For Tier 2 events the same summaries and confidence intervals will be provided, but no formal statistical testing. For Tier 3 events, estimates within each treatment group will be provided.

- Serious adverse experiences and patients withdrawn due to an adverse experience will be listed. Adverse experiences occurring before the first intake of study medication will be listed in the appendix of the report for disclosure. In addition, a listing will be provided in the appendix of the report of all occurrences of overdoses (ingesting 2 or more tablets of blinded study therapy on the same day or within a 5-day period).

In addition to the above mentioned summary tables, time to first event will also be summarized graphically by a Kaplan-Meier curve for any adverse experience, drug-related adverse experiences, serious adverse experiences, adverse experiences leading to early treatment discontinuation, deaths, and for any skin adverse experiences. Similar graphs will be provided for laboratory adverse experiences. For patients who did not experience an adverse experience of the specific category (any, related, serious, leading to discontinuation) before that time, the time to event will be censored at the cut-off date, or at the last contact date for patients who discontinued early from the trial before the cut-off date.

The main adverse experience tables (Tier 1 and 2 adverse experience summaries of Table 3-11) will also be provided, restricted to patients who are (oral) bisphosphonate-intolerant. Similar adverse experience summary tables will be provided for upper gastrointestinal adverse experiences, for patients who are (oral) bisphosphonate-intolerant.

Table 3-11

Analysis Strategy for Safety Parameters

Safety Tier	Safety Endpoint[†]	p-Value	95% CI for Treatment Comparison	Descriptive Statistics
Tier 1	Skin AEs reported as ECI[¶]	X	X	X
	Serious respiratory infections confirmed by adjudication[¶]	X	X	X
	Any osteonecrosis of the jaw confirmed by adjudication[¶]	X	X	X
	Any morphea/scleroderma confirmed by adjudication[¶]	X	X	X
	Any delayed fracture unions confirmed by adjudication[¶]	X	X	X
Tier 2	Any AE		X	X
	Any Serious AE		X	X
	Any drug-related AE		X	X
	Discontinuation due to AE		X	X
	Specific AEs[‡], Specific dental disorders, Specific skin disorders, Specific respiratory disorders (incidence ≥1% patients in one of the treatment groups), hypocalcemia		X	X
	PDLCs[‡] for lab tests (incidence ≥1% patients in one of the treatment groups)		X	X
Tier 3	Specific AEs[‡], specific dental disorders, specific skin disorders, specific respiratory disorders, PDLCs[‡] for lab tests (incidence > 0 patients in one of the treatment groups)			X
	Change from Baseline Results (Labs, Vital Signs)			X

[†] AE references refer to both Clinical and Laboratory AEs.
[‡] Includes only those endpoints not pre-specified as Tier 1 endpoints.
[¶] Only considered as Tier 1 if the incidence is ≥ 4 patients in one of the treatment groups, otherwise they will be considered Tier 3.
Note: PDLC=Pre-Defined Limit of Change.; X = results will be provided.

Change from baseline at each of the on-treatment time points in vital signs and laboratory safety endpoints will be summarized for both treatment groups. The number (%) of patients with laboratory values outside predefined limits of change will also be summarized similarly (with confidence intervals on the treatment difference). More detail is in Appendix 6.7.

Qualitative Abnormalities on Bone Biopsies

The 2-D Histomorphometry and 3-D μCT transilial bone biopsy measurements at Month 24 or the end of the study will be summarized as efficacy endpoints. The principle purpose of the histomorphometry and bone biopsies is to rule out an increased incidence of qualitative abnormalities (e.g. excessive osteoid, woven bone, mineralization defect, marrow fibrosis), the number (percent) of patients with a qualitative abnormality will therefore be summarized for both treatment groups, along with the Miettinen and Nurminen method [6] based 95% confidence interval on the difference in percentage. The incidence of qualitative abnormalities on the bone biopsies will be considered a safety endpoint in stead of an efficacy endpoint as described in the protocol.

3.5.5.3 Multiplicity

Several multiplicities are present in this study: multiple primary endpoints, multiple secondary endpoints, interim analyses. This section describes how the multiplicities for the final analysis will be handled. Multiplicity adjustment in the interims will be described in Section 3.5.5.6.

The study is planned to terminate and the final analysis will be performed when all pre-specified fracture endpoints have occurred. In addition, the DMC will review both efficacy (in the formal efficacy interims) and safety and can recommend early termination for either efficacy or safety, before all pre-specified events are seen, as described below.

Multiplicity Due to Multiple Primary Endpoints

This study is primarily designed to investigate the effect of odanacatib on morphometric vertebral, non-vertebral and hip fractures. To control the false positive error rate due to multiple fracture endpoints in the final analysis, a combination of a step-down closed-testing and Hochberg procedure [8] will be applied. First a step-down procedure will be utilized with the following order of clinical importance: (1) morphometric vertebral fractures, (2) hip and non-vertebral fractures.

This means that significance for hip and non-vertebral fractures can only be declared if there was also a statistically significant result for the morphometric vertebral fractures. If there is no statistical significance for morphometric vertebral fractures, no significant conclusions will be drawn from this study for any of the other primary endpoints.

To control the false positive error rate for multiple tests for hip and non-vertebral fractures, a Hochberg procedure [8] will be used. Under Hochberg procedure, p-values for treatment comparisons (for hip and non-vertebral fractures) will be ranked from the largest to the smallest (i.e. $P_1 > P_2$). If both p-values for hip and non-vertebral fractures are $\leq \alpha$, both corresponding tests will be considered significant and the testing stops. If however, the largest p-value is $>\alpha$, it is considered not significant and the other fracture endpoint will be tested at a $\alpha/2$ significance level.

A graphical presentation of the multiplicity adjustment strategy is in Figure 3-1 (Section 3.5.5.6).

In the interim analyses, the trial will not be terminated unless strong evidence (significance) is seen for all 3 primary fracture endpoints, more detail is provided below.

Multiplicity Adjustment for Secondary Endpoints

Secondary efficacy endpoints include clinical vertebral fractures, stature, BMD measures at different sites, and biochemical markers of bone resorption. Statistical significance for the secondary endpoints will only be considered if the treatment difference for the first primary endpoint (morphometric vertebral fractures) was significant. It is recognized that this does not account for the presence of multiple primary endpoints, but it is felt that this approach is clinically meaningful and acceptable.

For the purpose of addressing the issue of multiplicity due to multiple secondary endpoints, BMD measurements at the total hip, femoral neck, trochanter, lumbar spine, and distal forearm will be considered one family. Biochemical markers of bone resorption (s-CTX, u-NTx) will be considered a second family and clinical vertebral fractures and stature a third family. Within each family a Hochberg [8] multiplicity adjustment procedure will be used to ensure a family-wise type I error rate of 5% within each of the families (conditional on significance at later time points for BMD and height and conditional on significance at earlier time points for biomarkers). No adjustment for multiplicity across the families will be applied.

For BMD and height, a step-down procedure will be used to account for the multiplicity over time, starting from the last time point to the earliest time point where no significant conclusion can be claimed. At earlier time points significance will only be claimed if there was significance at the later time points and if from the Hochberg procedure including all other endpoints significance can be concluded, as explained in [24]. Since the very last timepoint in the study may be obtained only for a limited number of patients, the last timepoint which was performed for at least 50% of the patients will be taken as the last time point in this multiplicity adjustment strategy and the later time point(s) will be considered exploratory.

For biomarkers, a similar approach will be taken, starting from the earliest time point to the last time point where significance is claimed. This approach for biomarkers is taken because the treatment difference is not expected to increase over time after the first on-treatment time point (Month 6), while for BMD and height it is expected to increase.

Significance of the analyses of BMD endpoints in the subgroup of (oral) bisphosphonate intolerant patients will only be tested if the test for the same BMD endpoint in the full FAS population is significant after Hochberg adjustment procedure.

Alpha Adjustment for Interim Analyses

A DMC will monitor the safety of the patients in the study on a regular basis and will review efficacy in the formal efficacy interims. Before the first formal efficacy interim analysis, the DMC will mainly review safety and the study should not be terminated early for efficacy. To limit the effect of the interim analyses in this time frame on the final alpha level, a low alpha level will be used based on the same spending function used for the formal efficacy interim analyses. From the first formal efficacy interim onwards, the DMC will review both efficacy and safety and may recommend early termination of the trial both for efficacy or for safety. An alpha-spending function approach [25] will be used to handle the multiple looks due to these interim analyses. Details on the actual alpha levels are in Section 3.5.5.6.

3.5.5.4 Sample Size and Power Calculations

Underlying Assumptions

Sample size estimates are based on estimates of fracture incidence rates from the Fosamax Fracture Intervention Trial (FIT) data. It is assumed that the expected relative risk (RR) for morphometric vertebral fractures is 0.5, for non-vertebral fractures 0.8 and for hip fractures 0.65[2]. Sample sizes are based on a 4% alpha level at the final analysis (to account for the interim analyses, see multiplicity adjustment section, the final alpha will be ≥ 4%) and 90% power.

To ensure enrollment of a study population with adequate overall fracture risk it was decided to have a flexible sample size depending on the observed ratio of number of patients with a prior vertebral fracture to number of patients without prior vertebral fracture. It is expected that the total sample size needed will be smaller if more patients with a prior vertebral fracture can be included.

For the sample size calculations it is assumed that the recruitment period will be approximately 1 year and the total study duration (including the recruitment period) will be approximately 5 years. Other assumptions used to derive the different scenarios are based on estimates from the FIT data and are detailed in Table 3-12.

[2] In a recent study of the Novartis compound (ASBMR 28th annual meeting -sep 2006 - S16. Abstract #1054 D. Black et. al.) the relative risk for hip fractures was 0.6.

Table 3-12

Underlying Assumptions Used for Calculations of Sample Size, Based on FIT Data

Patient Population / Fracture Type	Approximate Placebo Fracture Rate	Relative Risk	Drop-out Rate
Patients with no prior vertebral fracture			
Vertebral fractures	4% at 2 years	0.5	10% with no data at all
Hip fractures	1.4% at 3 years	0.65	20% over 3 years
Non-vertebral fractures	8% at 2 years	0.8	13% over 2 years
Patients with 1 prior vertebral fracture			
Vertebral fractures	8% at 2 years	0.5	10% with no data at all
Hip fractures	3% at 3 years	0.65	20% over 3 years
Non-vertebral fractures	11% at 2 years	0.8	13% over 2 years

For the power and sample size calculations the morphometric vertebral fractures endpoint was considered to be binomially distributed at Month 24, the interval-censored survival approach including all patients data available is more powerful. For non-vertebral and hip fractures the primary focus is the number of events needed and time to event methodology was used, with exponential time-to-event and an exponential drop-out mechanism. The s-plus survival procedure using the Lakatos (or Rubinstein) method [26; 27] was used for the sample size and power calculations.

Although enrollment will not be uniform and there will be a Lead Cohort of 1,500 patients randomized earlier in time, sample size calculations are based on uniform enrollment over 1 year as an approximation. It is expected that the 1,500 patients in the Lead Cohort will be enrolled in approximately 3 - 4 months, and that the remaining 15,000 patients in the Main Cohort will be enrolled in approximately 8 months, after a gap of approximately 10 months with no recruitment. This means that the time between the first and last patients enrolled will be about 22 months, with 2 active enrollment periods adding up to a total of approximately 12 months of active recruitment. Due to the uncertainly about the actual enrollment distribution, and due to the fact that additional patient-exposure from the Lead Cohort will translate into a potentially earlier termination of the trial (when all events are seen), it is felt that the uniform distribution over 1 year with a 4 year study duration is a reasonable approximation.

Sample Size Calculations

We will attempt to enroll approximately 4,000 patients with one prior vertebral fracture into the trial. Since the fracture incidence is higher in this group of patients, a lower total number of patients can be included if more patients with a prior vertebral fracture are randomized. If, for example, approximately 3,400 patients with 1 prior vertebral fracture can be recruited in 12 months and 13,600 patients with no prior vertebral fracture (17,000 patients in total), the assumed hip fracture incidence in placebo is approximately

1.7 based on this ratio of patients with and without prior vertebral fracture, and there will be approximately 90% power (α=0.04) to detect a treatment difference of 35% reduction between the treatments for hip fractures, in favor of the odanacatib group. Other scenarios are detailed in Table 3-13.

Table 3-13

Approximate Sample Sizes for Different Ratios of Patients
(Those With and Those Without Prior Vertebral Fractures)
(Assuming 1 Year Recruitment, 4 Year Total Study, α=0.04, Power=90%)

Ratio of Patients With to Without Prior Fracture	Hip Fracture Rate in Placebo	Relative Risk	Drop-out Rate Over 3 Years	Number of Events	Approximate Total Number of Patients	Number of Patients Without Prior Fracture	Number of Patients With a Prior Fracture
Alpha = 0.04							
1:5	1.7	0.65	20%	237	17,600	14,650	2,950
1:4	1.7	0.65	20%	237	17,000	13,600	3,400
1:3	**1.8**	**0.65**	**20%**	**237**	**16,300**	**12,200**	**4,100**
1:2	1.9	0.65	20%	237	15,100	10,050	5,050
1:1	2.2	0.65	20%	236	13,250	6,625	6,625
2:1	2.5	0.65	20%	235	11,800	3,950	7,850

It is expected that approximately 1 out of 4 randomized patients will have a prior vertebral fracture. This means that a total of ~16,300 patients, 12,200 with no prior vertebral fracture and 4,100 with a prior vertebral fracture will have to be randomized. The size of the sample may be revised downward if a higher than anticipated proportion of patients with a prior fracture is enrolled. Conversely, it may be revised upward if a lower than expected proportion of such patients is enrolled. Any such revision will be done in consultation with the Protocol 018 Steering Committee. More detail is below.

Sensitivity Analyses

Table 3-14 provides sample size calculations for a 0.6 relative risk in both groups of patients. It can be seen that sample sizes are highly influenced by the assumed relative risk.

Table 3-14

Approximate Sample Sizes for Different Ratios of Patients
(Those with and Those without Prior Vertebral Fractures)
(Assuming 1 Year Recruitment, 4 Year Total Study, RR=0.60
in Both Groups, α=0.04, power=90%)

Ratio of Patients With to Without Fracture	Hip Fracture Rate in Placebo	Relative Risk	Drop-out Rate Over 3 Years	Number of Events	Approximate Total Number of Patients	Number of Patients Without Prior Fracture	Number of Patients With a Prior Fracture
Alpha = 0.04 - Relative Risk = 0.60 in both groups							
1:4	1.7	0.6	20%	169	12,750	10,200	2,550
1:3	1.8	0.6	20%	169	12,150	9,100	3,050
1:2	1.9	0.6	20%	168	11,300	7,550	3,750
1:1	2.2	0.6	20%	168	9,900	4,950	4,950

Power for Vertebral and Non-vertebral Fractures and BMD

All sample size calculations above are based on the estimates for the hip fracture incidence, since the analyses for morphometric vertebral and non-vertebral fractures have more power than those for hip fractures. For example, if 4,100 patients with one prior vertebral fracture and 12,200 patients without a prior vertebral fracture are randomized into the trial (see Table 3-13), there will be >99% power to detect a treatment difference in morphometric vertebral fractures and in non-vertebral fractures, using the assumptions of Table 3-12 and a significance level of 0.04.

With a total of at least 12,000 patients, there will be >99% power to detect a clinically meaningful BMD difference. For example, assuming that there are 6,000 patients per group, the common standard deviation (SD) is 4%[3], alpha is 5% and power of 99% the minimal detectable difference is 0.3%. When 1,200 (10%) patients have a BMD measurement (for distal forearm and total body), a 1% difference can be detected with 99% power (α=5%, SD=4%).

Note on Hochberg Procedure

As detailed in Section 3.5.5.3, a combination of a step-down and Hochberg procedure will be used to account for multiplicity due to multiple primary endpoints. This means that if the p-value for non-vertebral fractures is >α, hip fractures need to be tested at α/2. Similarly, if the p-value for hip fracture is >α, then non-vertebral fractures need to be

[3] The estimated 4% standard deviation was obtained from the phase IIb study and previous alendronate studies.

tested at $\alpha/2$. The actual alpha of the final analysis is detailed in Section 3.5.5.6 and will be between 4% and 5%.

To get some idea of the loss in power, the raw powers (not taking multiplicity adjustments or correlations between endpoints into account) for hip and non-vertebral fractures at 2% significance level are also calculated.

Assuming that approximately 16,300 patients (1:3 ratio of patients with to patients without prior vertebral fracture) are randomized and the underlying assumptions are as detailed earlier, then there is >99% power to detect a difference in non-vertebral fractures (RR=0.8) at 2% significance level. Similarly, the raw power for hip fractures is 84% to detect a treatment difference (RR=0.65) with α=2%. Note, as there is >99% to detect a treatment difference in non-vertebral fractures at α=4%, there is minimal chance that the test for hip fractures will need to be performed at α=2%, with the underlying assumptions as before.

Determination of Number of Patients to Enroll

As mentioned earlier, we estimate that a total of 237 hip fractures (first hip fracture per patient) will be needed based on the 1:3 ratio assumption of patients with to patients without a prior vertebral fracture. A total of ~16,300 patients (12,200 with no prior vertebral fracture and 4,100 with a prior vertebral fracture) will therefore have to be randomized. However, the SPONSOR (blinded to treatment assignment) will monitor this ratio regularly during the enrollment period, and calculate the targeted total number of patients to be randomized and the expected total study duration based on the observed ratio of patients with to patients without prior vertebral fracture, and the same assumptions of hip fracture rates as in Table 3-12 and Table 3-13, i.e., 1.4% hip fractures at 3 years for patients without prior vertebral fracture and 3% for those with prior vertebral fracture in the placebo group and an overall relative risk of 0.65, assuming the current recruitment trend continues. If needed, the SPONSOR can decide to stop recruitment in one of the strata, to allow the other stratum to accumulate or they may propose to prolong the enrollment period beyond 1 year, to achieve the necessary total number of patients. The total sample size will be decided upon based on the observed ratio of patients with a prior vertebral fracture to those without a prior vertebral fracture and the expected total study duration, by the SPONSOR.

The roles and responsibilities of the Steering Committee will be detailed in the Steering Committee charter.

Study Termination and Number of Events Needed for Hip, Vertebral and Non-vertebral Fractures

The total number of hip fractures needed to detect a treatment difference (RR=0.65) at 90% power with the underlying assumptions of Table 3-12 is approximately 237 events, the number of non-vertebral fractures is 824, and the number of morphometric vertebral fractures at Month 24 would be 114. Since the analyses are based on time to first fracture,

the above mentioned number of fractures needed are only the first events per patient. If a patient has, for example, 2 hip fractures, only the first hip fracture will be counted in the above numbers.

During the double-blind therapy period, the Steering Committee and SPONSOR will monitor the number of (first) hip, spine and non-vertebral fracture events observed (blinded) and determine when the study should be terminated. The DMC or the unblinded statistician will not be responsible for determining when the pre-specified number of events are seen, since they are unblinded to study data. Once the SPONSOR, based on the Steering Committee's recommendation, decides to terminate the study, all sites will be informed of this decision and all patients will be requested to come in for their close-out visit as soon as possible after this communication. Once all patients have completed this close-out visit, the database will be locked and unblinded and the final analysis performed.

The Steering Committee and/or Sponsor will also pre-specify a data cut-off date after the last expected close-out visit date. Any fractures or adverse events occurring after the data cut-off will not be included in the final analysis and patients without an event (and who were not discontinued from the trial before) will be censored at this cut-off date, as detailed in Section 3.5.4.

Power for the Analysis in the Subgroup of (Oral) Bisphosphonate Intolerant Patients

It is expected that approximately 1400 patients per group will be (oral) bisphosphonate intolerant and therefore included in the BMD analysis for the corresponding secondary hypothesis. With 1400 patients per group, there is >99% power to detect a 1% difference in percent change from baseline between Odanacatib 50 mg once weekly and placebo, assuming a ≥4.5% alpha level (after adjustment for the interim) and a 4% common standard deviation. With a 3% standard deviation this power is also >99%.

3.5.5.5 Effect of Baseline Factors and Subgroup Analyses

The primary and key secondary endpoints will be summarized by stratum (prior vertebral fracture, no prior vertebral fracture) and the consistency of the treatment effect across strata will be investigated as detailed in Section 3.5.5.1 by summary statistics per stratum.

In addition, for the primary endpoints (morphometric vertebral, non-vertebral and hip fractures), differential treatment effects will be explored across various subgroups, e.g., age, race, baseline BMD T-score tertiles, baseline biochemical marker tertiles, geographic region (e.g., Japan and Far East Region), prior vertebral fracture (0, 1, >1), bisphosphonate intolerance (yes/no), and use of intranasal calcitonin (yes/no). Percent change from baseline in BMD endpoints will also be summarized by bisphosphonate intolerance (yes/no) as well as for patients who are not candidates for or who refuse osteoporosis treatment with bisphosphonates or strontium or PTH versus the other patients. Other subgroups to be investigated are detailed in appendix 6.7. Fracture incidences will be provided for each of the subgroups (by treatment).

3.5.5.6 Interim Analyses

Procedures and Role of the Data Monitoring Committee

A Data Monitoring Committee (DMC) will monitor the safety of the patients during the conduct of the study. They will be responsible for ongoing surveillance of emerging safety data and will review unblinded safety information on a regular basis, e.g. approximately every 4 months. They will also be responsible for ensuring that the study data are analyzed appropriately and make recommendations to continue or terminate the trial early based on their review of the unblinded data either for efficacy or safety. As part of its regular safety reviews, the DMC will review WBC and differential counts unblinded and in aggregate.

DMC Safety Review before the First Formal Efficacy Interim (70% hip fracture events)

Before the first formal efficacy interim analysis (70% of planned hip fractures, as detailed below), the DMC will mainly review unblinded safety results. If they think efficacy results are necessary to make a proper risk-to-benefit assessment, these can be requested. To allow the study to accumulate sufficient longer-term safety information, the DMC should not recommend early termination for strong beneficial effects during this period.

DMC Safety Review in Formal Efficacy Interims (70% and 85% hip fracture events)

From the first formal efficacy interim analysis onwards, the DMC will review both efficacy and safety results and may recommend to continue or terminate the trial based on efficacy and/or safety data. If all three p-values (for vertebral, hip, non-vertebral fractures) are lower than the corresponding boundary based on the alpha-spending function at an interim look (see below), the DMC may recommend termination of the trial before all pre-specified events are seen. This formal stopping guideline is based on the primary endpoints. However, all available data, including secondary efficacy endpoints and safety results, should be considered in an aggregate fashion along with the results of the stopping rule to draw a conclusion to stop the trial.The DMC should specifically evaluate if sufficient long-term safety information is obtained at the time of the interim analyses before recommending termination of the trial.

Safety reviews of the data will coincide with the two efficacy interims. But as detailed in Section 3.5.5.3, additional interims may be added, or the 85% interim analysis may be dropped if deemed appropriate.

Blinding Procedures, Roles and Responsibilities of the DMC and Steering Committee

Based on its review of the unblinded results, the DMC will make recommendations to continue or terminate the trial before the total number of pre-specified events are seen. The SPONSOR, in consultation with the Steering Committee will then review this recommendation (without receiving unblinded study results) and will make a final decision with respect to the continuation or termination of the trial. No unblinded information should be disseminated outside the DMC, except for the instance in which

the DMC recommends early termination of the trial and wishes to show limited unblinded results to justify its conclusion. In this case, the number of people receiving this unblinded information should be kept to a minimum, the unblinding should be documented in a memo, and those unblinded should not be involved in the conduct of the trial, screening of data or determination of protocol violators from that point onward. Additional detail will be provided in the DMC charter. Once a final decision to terminate the trial is made and the database is locked for the final analyses, results can be distributed to a wider audience, according to Merck procedures for distribution of study results.

For the purposes of these periodic safety and efficacy reviews, a Merck statistician and/or statistical programmer (if needed) will be unblinded to the database, but will not be further involved in the conduct of the study, screening of data or determination of protocol violators, once unblinded.

The DMC can consult additional expert members to provide subject matter expertise. These external consultants will then adhere to the same rules as the DMC members with respect to confidentiality and should not be involved in the conduct of the study.

The roles and responsibilities of the DMC will be detailed in separate DMC charter.

Lead and Main Cohort Enrollment

As outlined earlier, the current trial will be enrolled in 2 separate phases. This two-phase approach to study enrollment is designed to avoid unnecessarily exposing large numbers of patients to odanacatib, and to permit study of a limited number of patients with close monitoring by an independent Data Monitoring Committee (DMC) to demonstrate that adverse experiences similar to those seen with balicatib are not associated with odanacatib treatment. In the first phase of enrollment, approximately 1,500 patients will be randomized ("Lead Cohort"). After approximately 1,500 patients have been enrolled, enrollment will stop until all patients in the 'Lead Cohort' have received study drug or placebo for at least 9 months. During this 9-month period, data will be reviewed by the DMC approximately every 4 months.

After all Lead Cohort patients have completed at least 9 months of treatment (or have discontinued from the trial early), the main safety tables to be used for the Final Analysis (Section 3.5.5.2), will be created and presented to the DMC by the unblinded statistician. The DMC will review this safety information and use standard signal detection methods (e.g., review of the 95% confidence intervals for treatment differences and use of clinical judgment) to make a risk/benefit assessment (considering efficacy data from the Phase IIb study). No efficacy analyses are planned for this safety review due to the limited number of patients included in the Lead Cohort. If the safety is found to be sufficiently reassuring by the DMC to continue, the second phase of recruitment will begin to enroll the balance of ~16,300 patients ('Main Cohort').

The Sponsor and relevant regulatory agencies will be informed of the DMC conclusion that the safety is sufficiently reassuring for enrollment of the main cohort to commence. No unblinded information will be shared outside the DMC.

More details on the actual tables to be provided to the DMC and specific criteria to determine when results are sufficiently reassuring will be determined by the protocol team and the DMC and documented in the DMC Charter.

Multiplicity Adjustment Strategy for the Interim Analyses

Before the First Formal Efficacy Interim (70% Hip Fracture Events)

Before the first formal efficacy interim analysis (when 70% of the pre-specified hip fracture events have occurred), the study should not be terminated for beneficial effects. A minimal alpha adjustment will therefore be made for the analyses during this period. The same alpha spending function will be used to calculate the cut-off values for the p-value, as is used for the 2 formal interim efficacy analyses (see below).

Formal Efficacy Interim Analyses

From the first formal efficacy interim onwards (see below), the DMC will also review efficacy results and recommend to continue or terminate the trial based on both efficacy and safety data. The first formal interim analysis for both efficacy and safety will be performed at the time at which approximately 70% of all pre-specified hip fracture events (first adjudicated osteoporotic hip fracture per patient) have occurred and is expected to be after approximately 3 years of total study duration (including 1 year recruitment). The second formal efficacy interim will be performed when approximately 85% of the pre-specified hip events have been observed, and the Final Analysis will occur when all pre-specified fracture events (hip, morphometric vertebral and non-vertebral) have been observed. Once the efficacy interim analyses begin, safety updates will be coordinated to be performed at the same time as these efficacy interims, and will no longer be regularly scheduled, e.g. every 6 months. But since the study is expected to be approximately 5 years in duration, and the first interim is expected to occur near the 3 year time point, the interims are expected to occur approximately every half-year as occurred in the first 3 years.

If the period between the first and second formal interim analyses or between the second formal interim and the Final Analysis is expected to be less than 3 months, the second formal interim analysis may be dropped and the alpha level will be recalculated accordingly. Similarly, if the period between these analyses is expected to be more than a year, additional interim analyses may be added to ensure sufficient safety review of the data, and the alpha levels will also be recalculated accordingly. The timing of the interim analyses is based on the proportion hip fracture events, which will be determined by the SPONSOR, in consultation with the Steering Committee, during their regular review of the blinded study information. In addition, if there is insufficient safety exposure to terminate the study even if an interim analysis demonstrates significant fracture risk

reduction, the SPONSOR, in consultation with the Steering Committee, may postpone the first and/or second efficacy interim. The corresponding alpha level will be recalculated in a blinded fashion.

The corresponding alpha for early termination for efficacy will be calculated using an alpha-spending function of the Gamma family with parameter -6 using EAST® software (Cytel software). The same alpha will also be used for the analyses of morphometric vertebral, hip and non-vertebral fractures in the first interim analysis. The study may be terminated for efficacy in the interim analyses if all 3 endpoints show significance at this calculated alpha level. The same approach will be used for the second interim analysis. The alpha level for efficacy will be calculated based on the percentage of hip fracture data with the same method and used for all 3 fracture endpoints.

For the final analysis, the final critical value for efficacy for each of the 3 endpoints will be calculated separately, using the EAST® alpha-spending function with user-specified boundaries, based on the observed amount of information per endpoint and the alpha spent in the 2 interim analyses. To be conservative, the minimum of these 3 calculated alpha cut-off values will then be used as the level for the final analysis and the multiplicity between the 3 endpoints will be handled by a combination of a step-down closed-testing and Hochberg procedure (morphometric vertebral first, then Hochberg between hip, and non-vertebral fractures), as detailed in the multiplicity adjustment for primary endpoints, Section 3.5.5.3. The final cut-off value for the alpha level will be calculated in a blinded manner by the SPONSOR and reviewed with the Steering Committee. The alpha levels for interim analyses at 70% and 85% hip fracture events are in Table 3-15. Depending on the number and timing of the safety interims (before 70% hip fracture events), the actual alpha-levels may be slightly different. The SPONSOR will calculate the alpha-levels for each interim blindly.

Stopping for Futility in the Interims

In the first formal interim analysis (70% hip fracture events), the cut-off value for futility will be calculated using EAST® by a gamma beta-spending function with parameter -20. The trial may be terminated in the first formal interim analysis if the p-value for vertebral fractures is higher than the corresponding cut-off for futility (Table 3-15). The same approach will be used in the second formal interim analysis, the cut-off for futility will be calculated by the same beta-spending function and the trial may be terminated in the second formal interim analysis if the p-value for morphometric vertebral fractures is higher than the corresponding cut-off for futility. The formal stopping guidelines for futility are based on the first primary efficacy endpoint, but all available information (efficacy and safety) should be taken into account before making a final decision.

It was decided to base the futility stopping guideline on the first primary endpoint (morphometric vertebral fractures), while the alpha-, beta-spending functions are based on the number of hip fracture events observed during the trial, for the following reasons: (1) hip fracture has the lowest expected incidence among the 3 types of fracture endpoints, (2) hip fracture reduction is the most difficult to achieve based on both its

lower incidence, as well as biology. The hip site contains a greater proportion of cortical bone compared to the spine, which has a greater proportion of trabecular bone. Cortical bone has different histology and micro architecture than trabecular bone and also has a slower turnover. (3) Morphometric vertebral fracture risk reduction, by contrast, is expected to be the most easily achieved and the study is highly powered for morphometric vertebral fractures. If this is not achieved, it is highly unlikely that reduction of non-vertebral or hip fracture will be achieved.

Examples of alpha levels for futility are also in Table 3-15.

Table 3-15

Direct Monitoring Using $\alpha(-6,t)$, $\beta(-20,t)$
(Two-sided test at α=0.050, α, β-spending function of Hwang, Shih, and DeCani)

	Stopping for Overwhelming Efficacy				Futility Stopping	
Cumulative Information†	Alpha at Each Look	Nominal Critical Point		Cumulative Alpha	Alpha at Each Look	Beta Spent
		Lower	Upper			
70%	0.008	-2.645	2.645	0.008	0.990	0.000
85%	0.018	-2.363	2.363	0.020	0.663	0.005
100%‡	0.046	-1.994	1.994	0.050	-	0.100

† Based on hip fracture events.
‡ In the final analysis the cut-off level (alpha at the final look) will be calculated for each endpoint separately based on their cumulative amount of data in the interim analyses, and the final alpha will be the minimum of these 3 cut-off levels and will therefore be slightly lower than 0.046.
Depending on the number and timing of the safety updates (before 70% hip fracture events), the actual alpha-levels may be slightly different.

In the formal efficacy interim analyses, the above mentioned stopping rules will be employed to potentially terminate the trial early for overwhelming evidence of efficacy or for futility. For efficacy, if all three p-values (for morphometric vertebral, hip, non-vertebral fractures) are lower than their corresponding boundaries via the alpha-spending function, the DMC may recommend termination of the trial before all pre-specified events are seen. For futility, the trial may be terminated early if the p-value for morphometric vertebral fractures is larger than the alpha-level calculated from the spending function. The formal stopping guidelines are based on the primary endpoints. However, all available data, including secondary efficacy endpoints and safety results should be considered in an aggregate fashion along with the results of the stopping rule to draw a decision to stop the trial.

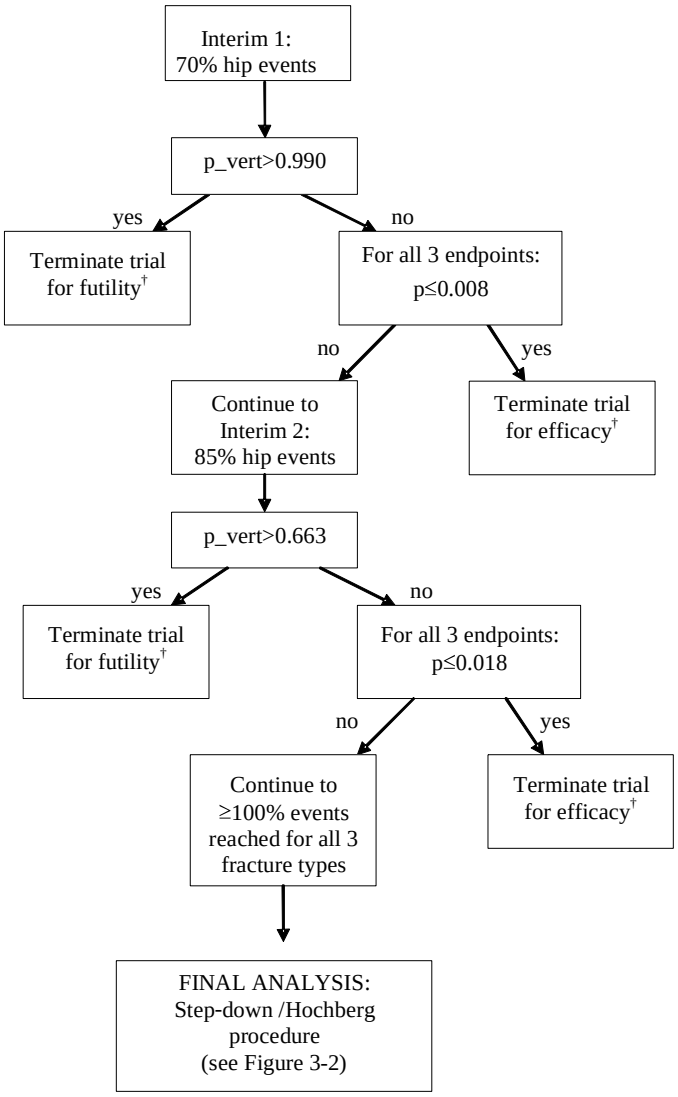
Summary of Interim Analyses and Multiplicity Adjustments

Figure 3-1 gives a graphical presentation of the decision process in the interim analyses as discussed above. Figure 3-2 shows the multiplicity adjustment in the final analysis, the actual cut-off value in the final analysis will be calculated as the minimum of the 3 cut-

off values for the 3 endpoints, as detailed above, and may be slightly lower than the presented 0.046.

Figure 3-1

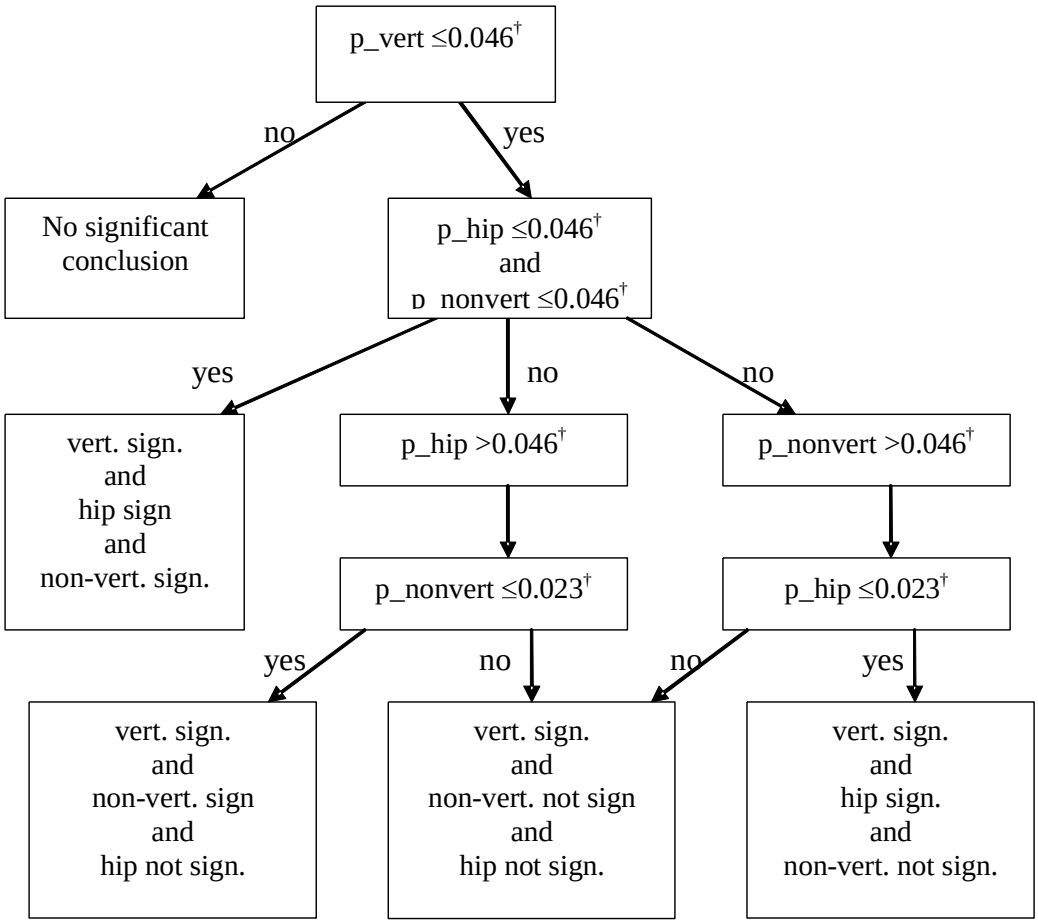
Interim Analyses



† All available efficacy and safety information should be taken into account to make a decision to terminate the trial.

Figure 3-2

Step-down and Hochberg Procedure in Final Analysis



† Final cut-off value to be calculated as the minimum of the 3 values for hip, vertebral fractures and non-vertebral fractures, using the alpha-spending functions and may be slightly smaller than 0.046.

3.5.6 Definition of Compliance Measure

For each patient, a compliance measure will be calculated based on the prime therapy records provided in the electronic Case Report Forms (CRF). Compliance will be defined as the percentage of the actual number of days with a double-blind treatment intake over the expected number of days with a treatment intake.

3.5.7 Extent of Exposure

Extent of exposure (dose and duration) will be summarized by treatment, more details are in appendix 6.7. Treatment duration will also be summarized for the lead and main cohorts separately.

3.6 LABELING, PACKAGING, STORAGE, DISPENSING, AND RETURN OF CLINICAL SUPPLIES

For studies using Controlled Substances, all Federal, State, Province, Country, etc., regulations must be adhered to in regard to the shipping, storage, handling, and dispensing of controlled substances. Additionally, the investigator should have the appropriate controlled drug license(s) as mandated by Federal, State, Province, Country, etc. laws in which the study is being conducted.

Investigational clinical supplies must be received by a designated person at the study site, handled and stored safely and properly, and kept in a secured location to which only the investigator and designated assistants have access. Clinical supplies are to be dispensed only in accordance with the protocol. The investigator is responsible for keeping accurate records of the clinical supplies received from the SPONSOR, the amount dispensed to and returned by the subjects/patients, and the amount remaining at the conclusion of the study. In accordance with Good Pharmacy Practices, gloves should always be worn by study personnel if directly handling tablets or capsules that are returned (i.e., when counting returns). The Clinical Monitor should be contacted with any questions concerning investigational products where special or protective handling is indicated. At the end of the study, all clinical supplies including partial and empty containers must be returned as indicated on the Contact Information page(s). U.S. sites should follow instructions for the Clinical Supplies Return Form (R464) and contact your SPONSOR representative for review of shipment and form before shipping. Sites outside of the United States should check with local country Merck personnel for appropriate documentation that needs to be completed for drug accountability.

3.6.1 Patient and Replacements Information

Clinical supplies will be packaged to support enrollment of approximately 19000 patients/subjects. Clinical supplies will be packaged according to a component schedule generated by the SPONSOR.

3.6.2 Product Descriptions

Investigational materials will be provided by the SPONSOR as summarized in Table 3-16.

Table 3-16

Product Descriptions

Product Name & Potency	Dosage Form	Comments
Odanacatib 50 mg	Tablet	Clinical image
Placebo to match Odanacatib 50 mg	Tablet	Clinical image
Cholecalciferol (Vitamin D_3) 2800 IU	Tablet	Clinical image
Calcium carbonate 500 mg	Tablet	Market image – for US and Japanese sites only

Calcium carbonate 1250 mg (500 elemental calcium) tablets will be supplied to the U.S. and Japanese study sites by the SPONSOR [see Section 6.6]; a supplement with similar calcium content, as approved by the local clinical monitor, will be provided by the investigational sites in all other countries. Patients will be provided with supplemental calcium such that their daily calcium intake (from diet and supplement) is approximately 1200 mg, but not to exceed 1600 mg. The investigator or designee will record the lot number, expiration date, and the quantity of drug dispensed. Alternative calcium supplements for U.S. and Japanese patients will only be allowed with the prior approval of the respective clinical monitor. However, Merck & Co., Inc. will neither provide nor reimburse these. In the event that the Merck-supplied Vitamin D_3 is not available at the site, an alternative supplement may be used as approved by the SPONSOR.

3.6.3 Primary Packaging and Labeling Information

Blinded supplies will be packaged in HDPE bottles as described in Table 3-17 below.

Table 3-17

Packaging of Blinded Clinical Supplies

Interval ID	Product Name & Potency	Fill Count	Dosing Instructions
Visit (blank space to be completed by investigator or designee)	Odanacatib 50 mg or matching placebo	16 tablets	Every week, take 1 tablet on your chosen day.

Container label text may include the following:

Packaging Control #/Packaging Lot trace ID #	Dosing Instructions
Space for baseline #	Storage Conditions
Component ID #	Compound ID - Protocol #
Space for Allocation #	Country regulatory requirements
Fill Count & Dosage Form	SPONSOR address (If applicable)
Interval ID	Translation Key (If applicable)
Re-evaluation date	

Open label supplies will be packaged in HDPE bottles as described in Table 3-18 below. Vitamin D_3 supplies will be packaged in brown amber bottles. A sample calcium label is attached as Appendix 6.6.

Table 3-18

Packaging of Open Label Clinical Supplies

Interval ID	Product Name & Potency	Fill Count	Dosing Instructions
Visit (blank space to be completed by investigator or designee)	Cholecalciferol (Vitamin D_3) 2800 IU	32 tablets	Every week, take 2 tablets on your chosen day.
Visit (blank space to be completed by investigator or designee)	Calcium carbonate 500 mg	250 tablets	Take (BLANK) tablet(s) daily, preferably with a meal.

Container label text may include the following:

Packaging Control #/Packaging Lot trace ID # Component ID# Space for baseline # Space for Allocation # Fill Count & Dosage Form Interval ID Product name and potency Re-evaluation date	Dosing Instructions Storage Conditions Compound ID - Protocol # Country regulatory requirements SPONSOR address (If applicable) Translation Key (If applicable)

3.6.4 Secondary Packaging and Labeling Information (Kit)

Supplies will NOT be packaged in kit boxes.

3.6.5 Clinical Supplies Disclosure

The IVRS should be used in order to unblind patients and to unmask drug identity. The SPONSOR will not provide disclosure envelope with the clinical supplies. Drug identification information is to be unmasked ONLY if necessary for the welfare of the patient. Every effort should be made not to unblind the patient unless necessary. Prior to unblinding, the investigator will attempt to contact the Clinical Monitor; however the Investigator may unblind a patient for safety reasons without first contacting the Clinical Monitor. Any unblinding that occurs at the site must be documented.

3.6.6 Storage and Handling Requirements

Clinical supplies should be kept in a secured location. Odanacatib supplies should not be stored above 30ºC. Calcium carbonate supplies should be stored between 15-30ºC – to

preserve quality and freshness, keep bottle tightly closed. Cholecalciferol (Vitamin D_3) supplies should not be stored above 25ºC and protected from light and moisture.

The clinical supplies storage area at the site must be monitored by the site staff for temperature consistency with the acceptable storage temperature range specified in this protocol or in the product label attached to the protocol. Documentation of temperature monitoring should be maintained.

3.6.7 Standard Policies/Return of Clinical Supplies

Investigational clinical supplies must be received by a designated person at the study site, handled and stored safely and properly, and kept in a secured location to which only the investigator and designated assistants have access. Clinical supplies are to be dispensed only in accordance with the protocol. The investigator is responsible for keeping accurate records of the clinical supplies received from the SPONSOR, the amount dispensed to and returned by the patients, and the amount remaining at the conclusion of the study. In accordance with Good Pharmacy Practices, gloves should always be worn by study personnel if directly handling tablets or capsules that are returned (i.e., when counting returns). The Clinical Monitor should be contacted with any questions concerning investigational products where special or protective handling is indicated. At the end of the study, all clinical supplies including partial and empty containers must be returned as indicated on the Contact Information page(s). U.S. sites should follow instructions for the Clinical Supplies Return Form (R464) and contact your SPONSOR representative for review of shipment and form before shipping. Sites outside of the United States should check with local country Merck personnel for appropriate documentation that needs to be completed for drug accountability.

The investigator or designated assistant should not open individual clinical supply containers and count tablets/capsules, etc., before dispensing to the patients. Any deviation from this must be discussed with the Clinical Monitor.

3.6.8 Distributing to Sites and Dispensing to Patients

Study personnel will have access to an Interactive Voice Response System (IVRS) to allocate patients, to assign drug to patients and to manage the distribution of clinical supplies. Each person accessing the IVRS system must be assigned an individual unique PIN. They must use only their assigned PIN to access the system and they must not share their assigned PIN with anyone.

3.7 DATA MANAGEMENT

Information regarding Data Management procedures for this protocol will be provided by the SPONSOR.

3.8 BIOLOGICAL SPECIMENS

All biological specimens should be collected, handled, stored and shipped according to the instruction provided in the laboratory manual.

4. ADMINISTRATIVE AND REGULATORY DETAILS

4.1 CONFIDENTIALITY

4.1.1 Confidentiality of Data

For Studies Conducted Under the U.S. IND

Particular attention is drawn to the regulations promulgated by the Food and Drug Administration under the Freedom of Information Act providing, in part, that information furnished to clinical investigators and Institutional Review Boards will be kept confidential by the Food and Drug Administration only if maintained in confidence by the clinical investigator and Institutional Review Board.

For All Studies

By signing this protocol, the investigator affirms to the SPONSOR that information furnished to the investigator by the SPONSOR will be maintained in confidence and such information will be divulged to the Institutional Review Board, Ethics Review Committee, or similar or expert committee; affiliated institution; and employees only under an appropriate understanding of confidentiality with such board or committee, affiliated institution and employees. Data generated by this study will be considered confidential by the investigator, except to the extent that it is included in a publication as provided in the Publications section of this protocol.

4.1.2 Confidentiality of Subject/Patient Records

For All Studies

By signing this protocol, the investigator agrees that the SPONSOR (or SPONSOR representative), Institutional Review Board/Independent Ethics Committee (IRB/IEC), or Regulatory Agency representatives may consult and/or copy study documents in order to verify worksheet/case report form data. By signing the consent form, the subject/patient agrees to this process. If study documents will be photocopied during the process of verifying worksheet/case report form information, the subject/patient will be identified by unique code only; full names/initials will be masked prior to transmission to the SPONSOR.

For Studies Conducted Under the U.S. IND

By signing this protocol, the investigator agrees to treat all patient data used and disclosed in connection with this study in accordance with all applicable privacy laws, rules and regulations, including all applicable provisions of the Health Insurance Portability and Accountability Act and its implementing regulations, as amended from time to time ("HIPAA").

4.1.3 Confidentiality of Investigator Information

For All Studies

By signing this protocol, the investigator recognizes that certain personal identifying information with respect to the investigator, and all subinvestigators and study site

personnel, may be used and disclosed for study management purposes, as part of a regulatory submissions, and as required by law. This information may include:

- name, address, telephone number, and email address;
- hospital or clinic address and telephone number;
- curriculum vitae or other summary of qualifications and credentials; and
- other professional documentation.

Consistent with the purposes described above, this information may be transmitted to the SPONSOR, and subsidiaries, affiliates and agents of the SPONSOR, in your country and other countries, including countries that do not have laws protecting such information. Additionally, the investigator's name and business contact information may be included when reporting certain serious adverse events to regulatory agencies or to other investigators. By signing this protocol, the investigator expressly consents to these uses and disclosures.

For Multicenter Studies
In order to facilitate contact between investigators, the SPONSOR may share an investigator's name and contact information with other participating investigators upon request.

4.2 COMPLIANCE WITH LAW, AUDIT, AND DEBARMENT

By signing this protocol, the investigator agrees to conduct the study in an efficient and diligent manner and in conformance with this protocol; generally accepted standards of Good Clinical Practice; and all applicable federal, state, and local laws, rules and regulations relating to the conduct of the clinical study.

The Code of Conduct, a collection of goals and considerations that govern the ethical and scientific conduct of clinical investigations sponsored by Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., is attached.

The investigator also agrees to allow monitoring, audits, Institutional Review Board/Independent Ethics Committee review, and regulatory agency inspection of trial-related documents and procedures and provide for direct access to all study-related source data and documents.

The investigator agrees not to seek reimbursement from subjects/patients, their insurance providers, or from government programs for procedures included as part of the study reimbursed to the investigator by the SPONSOR.

The Investigator shall prepare and maintain complete and accurate study documentation in compliance with Good Clinical Practice standards and applicable federal, state, and local laws, rules and regulations; and, for each subject/patient participating in the study,

provide all data, and upon completion or termination of the clinical study submit any other reports to the SPONSOR as required by this protocol or as otherwise required pursuant to any agreement with the SPONSOR.

Study documentation will be promptly and fully disclosed to the SPONSOR by the investigator upon request and also shall be made available at the investigator's site upon request for inspection, copying, review, and audit at reasonable times by representatives of the SPONSOR or any regulatory agencies. The investigator agrees to promptly take any reasonable steps that are requested by the SPONSOR as a result of an audit to cure deficiencies in the study documentation and worksheets/case report forms.

International Conference of Harmonization Good Clinical Practice guidelines (Section 4.3.3) recommend that the investigator inform the subject's primary physician about the subject's participation in the trial if the subject has a primary physician and if the subject agrees to the primary physician being informed.

According to European legislation, a SPONSOR must designate a principal or coordinating investigator (CI) to review the report (summarizing the study results) and confirm that to the best of his/her knowledge the report accurately describes conduct and results of the study. The SPONSOR may consider one or more factors in the selection of the individual to serve as the CI (e.g., thorough understanding of clinical trial methods, appropriate enrollment of subject/patient cohort, timely achievement of study milestones, availability of the CI during the anticipated review process).

The investigator will promptly inform the SPONSOR of any regulatory agency inspection conducted for this study.

Persons debarred from conducting or working on clinical studies by any court or regulatory agency will not be allowed to conduct or work on this SPONSOR's studies. The investigator will immediately disclose in writing to the SPONSOR if any person who is involved in conducting the study is debarred, or if any proceeding for debarment is pending or, to the best of the investigator's knowledge, threatened.

In the event the SPONSOR prematurely terminates a particular trial site, the SPONSOR will promptly notify that site's IRB/IEC.

4.3 COMPLIANCE WITH FINANCIAL DISCLOSURE REQUIREMENTS

By signing this protocol, the investigator agrees to provide to the SPONSOR accurate financial information to allow the SPONSOR to submit complete and accurate certification and disclosure statements as required by U.S. Food and Drug Administration regulations (21 CFR Part 54). The investigator further agrees to provide this information on a Financial Disclosure/Certification Form that is provided by Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc. This requirement also extends to subinvestigators. The investigator also consents to the transmission of this information to Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., in the United States for

these purposes. This may involve the transmission of information to countries that do not have laws protecting personal data.

4.4 QUALITY CONTROL AND QUALITY ASSURANCE

By signing this protocol, the SPONSOR agrees to be responsible for implementing and maintaining quality control and quality assurance systems with written SOPs to ensure that trials are conducted and data are generated, documented, and reported in compliance with the protocol, accepted standards of Good Clinical Practice, and all applicable federal, state, and local laws, rules and regulations relating to the conduct of the clinical study.

4.5 COMPLIANCE WITH INFORMATION PROGRAM ON CLINICAL TRIALS FOR SERIOUS OR LIFE THREATENING CONDITIONS

Under the terms of the Food and Drug Administration Modernization Act (FDAMA), the SPONSOR of the study is solely responsible for determining whether the study is subject to the requirements for submission to the Clinical Trials Data Bank. In accordance with the criteria set forth in the FDA "Guidance for Industry: Information Program on Clinical Trials for Serious or Life Threatening Conditions," Mar-2002, Merck, as SPONSOR of this study, has reviewed this protocol, determined that it meets the criteria for submission to the Clinical Trials Data Bank, and will submit the information necessary to fulfill this requirement.

By signing this protocol, the investigator acknowledges that the statutory obligation under FDAMA is that of the SPONSOR and agrees not to submit any information about this study to the Clinical Trials Data Bank.

4.6 PUBLICATIONS

As this study is part of a multicenter trial, publications derived from this study should include input from the investigator(s) and SPONSOR personnel. Such input should be reflected in publication authorship, and whenever possible, preliminary agreement regarding the strategy for order of authors' names should be established before conducting the study. Subsequent to the multicenter publication, or 24 months after completion of the study, whichever comes first, an investigator and/or his/her colleagues may publish the results for their study site independently. However, the SPONSOR does not recommend separate publication of individual study site results due to scientific concerns.

The SPONSOR must have the opportunity to review all proposed abstracts, manuscripts, or presentations regarding this study 60 days prior to submission for publication/presentation. Any information identified by the SPONSOR as confidential must be deleted prior to submission. SPONSOR review can be expedited to meet publication guidelines.

5. LIST OF REFERENCES

1. Cooper C. Epidemiology of osteoporosis. Osteoporos Int 1999;2(Suppl 2):S2-S8.

2. Melton LJ, III, Chrischilles EA, Cooper C, Lane AW, Riggs BL. How many women have osteoporosis? J Bone Miner Res 1992;7(9):1005-10.

3. National Osteoporosis Foundation: Fast Facts.

4. Wiktorowicz ME, Goeree R, Papaioannou A, Adachi JD, Papadimitropoulos E. Economic implications of hip fracture: health service use, institutional care and cost in Canada. Osteoporos Int 2001;12:271-8.

5. Rünger TM, Adami S, Benhamou CL, Czerwinski E, et al. Morphea-like skin reactions in patients treated with the cathepsin K inhibitor balicatib. J Am Acad Dermatol 2010.11.033.

6. Lutgens E, Lutgens SPM, Faber BCG, Heeneman S, Gijbels MMJ, de Winter MPJ, et al. Disruption of the cathepsin K gene reduces atherosclerosis progression and induces plaque fibrosis but accelerates macrophage foam cell formation. Circulation 2006;113:98-107.

7. Lutgens SPM, Kisters N, Lutgens E, van Haaften RIM, Evelo CTA, de Winther MPJ, et al. Gene profiling of cathepsin K deficiency in atherogenesis: profibrotic but lipogenic. J Pathol 2006;210:334-43.

8. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 1988;75(4):800-2.

9. Levine RJ. Placebo controls in clinical trials of new therapies for osteoporosis. J Bone Miner Res 2003;18(3):1154-9.

10. Ragi-Eis S, Zerbini CAF, Provenza JR, Griz LHM, de Gregório LH, Russo LAT, et al. Is it ethical to use placebos in osteoporosis trials? J Clinical Densitometry 2006; 9(3):274-80.

11. Panneman MJM, Lips P, Sen SS, Herings RMC. Undertreatment with anti-osteoporotic drugs after hospitalization for fracture. Osteoporos Int 2004;15:120-4.

12. Cramer, J. A., Amonkar, M. M., Hebborn, A., and Suppapanya, N. Does dosing regimen impact persistence with bisphosphonate therapy among postmenopausal osteoporotic women?

13. Chapuy MC, Arlot ME, Delmans PD, Meunier PJ. Effect of calcium and cholecalciferol treatment for three years on hip fractures in elderly women [Abstract]. BMJ 1994;308:1081-2.

14. Trivedi DP, Doll R, Khaw KT. Effect of four monthly oral vitamin D_3 (cholecalciferol) supplementation on fractures and mortality in men and women living in the community: randomised double blind controlled trial. BMJ 2003;326:1-6.

15. Bischoff-Ferrari HA, Willett WC, Wong JB, Giovannucci E, Dietrich T, Dawson-Hughes B. Fracture prevention with vitamin D supplementation. JAMA 2005; 293(18):2257-64.

16. Kanis JA, Johnell O, De Laet C, Johansson H, Oden A, Delmas P, et al. A meta-analysis of previous fracture and subsequent fracture risk. Bone 2004;35:375-82.

17. De Laet CE, van Hout BA, Burger H, Hofman A, Pols HA. Bone density and risk of hip fracture in men and women: cross sectional analysis [Abstract]. BMJ 1997; 315:221.

18. Rosen CJ, Hochberg MC, Bonnick SL, McClung M, Miller P, Broy S, et al. Treatment with once-weekly alendronate 70 mg compared with once-weekly risedronate 35 mg in women with postmenopausal osteoporosis: a randomized double-blind study. J Bone Miner Res 2005;20(1):141-51.

19. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst 2006;98(17):1183-92.

20. Potti A, Mukherjee S, Petersen R, Dressman HK, Bild A, Koontz J, et al. A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. N Engl J Med 2006;355(6):570-80.

21. Collet D. Interval-censored survival data. In: Chapman and All, ed. Modeling Survival Data in Medical Research. 1994:237-51.

22. Whitehead J. The analysis of relapse clinical trials, with application to a comparison of two ulcer treatments. Stat Med 1989;8:1439-54.

23. Miettinen O, Nurminen M. Comparative analysis of two rates*. Stat Med 1985;4:213-26.

24. Quan H, Luo X, Capizzi T. Multiplicity adjustment for multiple endpoints in clinical trials with multiple doses of an active treatment. Stat Med 2005;24:2151-70.

25. Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika 1983;70(3):659-63.

26. Lakatos E, Lan KKG. A comparison of sample size methods for the logrank statistic. Stat Med 1992;11:179-91.

27. Rubinstein LV, Gail MH, Santner TJ. Planning the duration of a comparative clinical trial with loss to follow-up and a period of continued observation. J Chron Dis 1981; 34:469-79.

28. Richards JB, Rivadeneira F, Inouye M, Pastinen TM, Soranzo N, Wilson SG, et al. Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. Lancet 2008;371:1505-12.

29. Mencej-Bedrac S, Preželj J, Kocjan T, Teskac K, Ostanek B, Smelcer M, et al. The combinations of polymorphisms in vitamin D receptor, osteoprotegerin and tumour necrosis factor superfamily member 11 genes are associated with bone mineral density. J Mol Endocrinol 2009;42:239-47.

30. Website: www.affymetrix.com, as of 01-Aug-2009.

31. Illumina sequencing technology, 2009.

6. APPENDICES

6.1 LABORATORY SAFETY ASSESSMENTS

Hematology

Blood hemoglobin
Whole blood hematocrit
Blood platelet count
White blood cell count (total and differential)

Blood Chemistry

Serum sodium
Serum potassium
Serum chloride
Serum blood urea nitrogen
Serum creatinine
Serum glucose
Serum aspartate aminotransferase
Serum alanine aminotransferase
Serum alkaline phosphatase
Serum calcium
Total serum protein
Serum albumin
Serum phosphorus
Serum magnesium
Total serum bilirubin
Direct serum bilirubin
Indirect serum bilirubin

Lipid profile and C-reactive protein measured at Randomization.

PTH measured at Screening in patients with a documented history of parathyroid disease

PTH and 25-hydroxyvitamin D measured at Screening in patients with a documented history of renal stones.

PTH and 25-hydroxyvitamin D measured at Screening in patients taking anti-seizure medication.

TSH measured at Screening in patients with a documented history of thyroid disease.

PTH, 25-hydroxyvitamin D, and TSH measured at Screening in cases where this is required by regulatory agency via documented request, subsequently approved by the SPONSOR.

PTH and 25-hydroxyvitamin D measured at baseline before Randomization in patients who at Screening have both serum creatinine >1.6 mg/dL and calculated creatinine clearance 30 – 59 mL/min.

Urinalysis

Dipstick in all patients; microscopic analysis performed only if dipstick abnormal

6.2 LABORATORY EFFICACY ASSESSMENTS

Performed in a random 10% subset of patients

s-BSAP (Bone specific alkaline phosphatase): Rand, Month 6, annually thereafter, end of study

s-P1NP (serum N-terminal propeptide of Type I collagen): Rand, Month 6, annually thereafter, end of study

u-NTx (N-telopeptides of Type 1 collagen): Rand, Month 6, annually thereafter, end of study

s-CTx (C-telopeptides of Type I collagen): Rand, Month 6, annually thereafter, end of study

PTH: Rand, Month 6, annually thereafter, end of study

25-hydroxyvitamin D: Rand, Month 12, end of study

Archives at each timepoint in the 10% subset; Archives for *all* patients at Randomization, Month 12, and end of study

PK analysis

Plasma for possible PK analysis at Randomization, and Months 3, 6, and 9 for the first 1500 patients enrolled.

6.3 APPROXIMATE RADIATION EXPOSURES AND BLOOD DRAW VOLUMES

Below is a listing of the approximate radiation exposure associated with the procedures in this protocol:

Radiation Exposures Associated with Procedures in Protocol

BMD	(Radiation Exposure in μSv)	
Hip [and subregions]	<10	Screening, annually, end of study
Spine[†]	<10	Randomization, annually, end of study
Forearm	<5	Screening, annually, end of study (10% subset)
Total body	<10	Screening, annually, end of study (10% subset)
Lateral Spine Radiograph	700	Screening, Month 6, annually, end of study

[†] Spine DXA at screening may only be performed in cases where this is required by regulatory agency via documented request, subsequently approved by SPONSOR.

Below is a listing of approximate blood volumes associated with the blood draws in this protocol. (Note: blood volumes may be decreased from the below if assay allows):

Approximate Blood Draw Volumes During the Study[†]

	Approximate Total Amount Per Collection
Chemistry, Lipid Profile, C-reactive Protein, and TSH[‡]	20 mL
Hematology	4 mL
PTH, 25-Hydroxyvitamin D, and Bone Biochemical Markers[‡]	20 mL
Archive[‡]	20 mL
Proteomics	10 mL
PK[‡]	4 mL
Genomics[§]	4 mL

[†] Note that because (a) this is an event-driven trial, and (b) patients will be enrolled in 2 phases, the total number of blood-draws will differ from patient to patient throughout the study.
[‡] Not all patients will have all blood draws at all visits, or require all tests indicated above. Please see Study Flow-Chart.
[§] Sample volume has been reduced from 10 mL to 4 mL in Protocol 018-02. Any genomics sample drawn prior to IRB/ERC approval and implementation of Protocol 018-02 is 10 mL in volume.

6.4 COMMITTEE STRUCTURE AND RESPONSIBILITIES

A. Data Monitoring Committee

The study will be conducted under the auspices of the DMC. The voting members of the committee will not include employees of the SPONSOR, and will not have any relation to study patients (i.e. not a study Investigator). The DMC will be appointed by the SPONSOR and will include 4 or 5 physicians and at least 1 statistician. A nonvoting, unblinded statistician employed by the SPONSOR will also attend DMC meetings.

The DMC will review data from the study at regular intervals and determine if it is safe to continue the study according to the protocol. The DMC will have access to all available data from the study throughout the study duration. Before the first formal efficacy interim analysis, the trial will not be stopped early for efficacy. (However, the DMC can request efficacy results for risk-benefit assessment, just not for early stopping conclusions). From the first formal efficacy analysis onwards, the DMC will review both efficacy and safety, as outlined in the Data Analysis Section of the protocol. The DMC will forward recommendations regarding proposed changes to the protocol, including interruption or discontinuation of the study to the SPONSOR.

Specific details regarding responsibilities and governance of the DMC will be described in a separate charter.

B. Fracture Trial Steering Committee

This study will be conducted in consultation with a Steering Committee, which will consist of SPONSOR personnel and Investigators participating in the trial and consulting osteoporosis experts and clinical trialists. This committee will provide guidance on the operational aspects of this trial. Specific details regarding responsibilities and governance of the Steering Committee will be described in a separate charter.

C. Endpoint Adjudication Committees

Fractures and Delayed Fracture Union AEs

All clinical fracture events (both vertebral and non-vertebral), with the exception of those of the fingers, toes, and face, will be evaluated by a Central Adjudication Committee (CAC). As is the case in all fracture endpoint trials, a determination will be made for each incident clinical fracture as to whether it is osteoporotic, traumatic or due to another cause.

Cases of possible delayed fracture union (fractures in which radiographic evidence of union is not present within 3 months [or within 6 months for tibial and femoral shaft

fractures] after the original fracture event) will be evaluated by an external fracture CAC.

Dental AEs

All dental AEs and dental procedures (other than routine cleaning) are to be reported. Suspected cases of "osteonecrosis of the jaw" (ONJ) evaluated by an external dental CAC. The investigators should pay special attention to the ONJ diagnosis criteria, especially delayed wound healing longer than 8 weeks. Suspected cases of ONJ must be followed until resolution.

Skin AEs

Skin AEs with skin thickening and hardening suggestive of morphea or systemic sclerosis will be evaluated by an external skin CAC for the presence or absence of morphea-like features.

Respiratory AEs

All respiratory adverse experiences meeting the regulatory definition of serious, except when lung cancer is the only diagnosis, will be evaluated by an external respiratory CAC.

Cardiovascular AEs

Cardiovascular (CV) events in the categories of thrombotic CV events (including acute and silent myocardial infarction, unstable angina pectoris, and cardiac thrombus), cardiac arrest, cardiac death, and sudden or unexplained death, and new onset atrial fibrillation and atrial flutter will be evaluated by an external cardiovascular CAC.

Cerebrovascular AEs

Cerebrovascular events in the categories of ischemic or hemorrhagic strokes and strokes of unknown mechanism will be evaluated by an external cerebrovascular CAC.

All personnel involved in the adjudication process will remain blinded to treatment allocation throughout the study. The details of the adjudication processes will be included in separate charters.

6.5 SAMPLE CALCIUM QUESTIONNAIRE

Dietary Calcium Questionnaire
From Dairy and Non-Dairy Sources

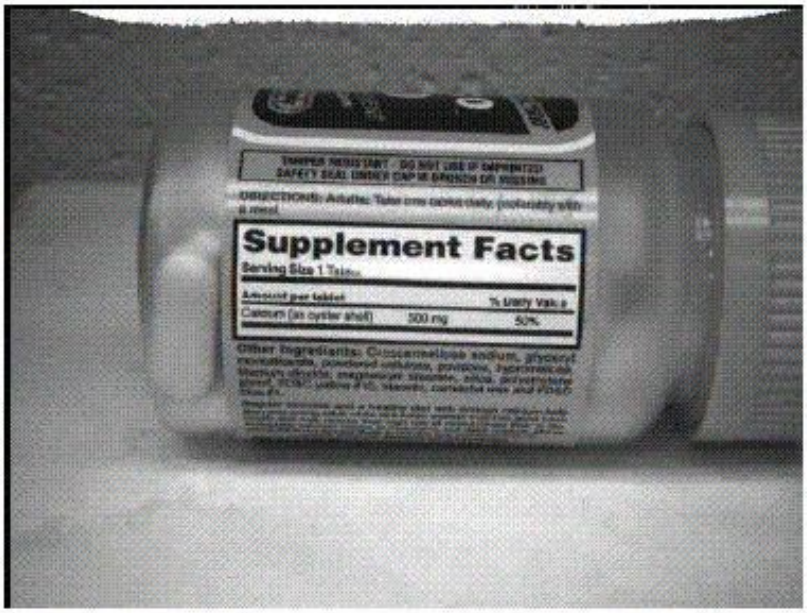
Food Type	Standard Serving Size	Usual Serving Size	Servings Per Week	Standard Servings Per Week	Calcium per Standard Serving (mg)	Calcium Per Week (mg)†
1. Milk (any type)	1 cup			x	292	=
2. Ice cream or other frozen desserts	1 cup			x	200	=
3. Hard cheese	1" cube or 1-oz slice			x	210	=
4. Grated parmesan cheese	1 Tb.			x	68	=
5. Cottage cheese	1 cup			x	212	=
6. Yogurt	1 cup			x	290	=
7. Half & Half	1 Tb.			x	16	=
8. Almonds	½ cup			x	200	=
9. Baked beans, soybeans, white beans	1 cup			x	150	=
10. Bok choy or kale, cooked	½ cup			x	75	=
11. Bread, whole wheat or white	1 slice			x	25	=
12. Broccoli, cooked	¾ cup			x	50	=
13. Chickpeas	1 cup			x	75	=
14. Kidney beans, lima beans, lentils	1 cup			x	50	=
15. Nuts, mixed	½ cup			x	48	=
16. Orange (fruit, not juice)	1 medium			x	50	=
17. Sardines, canned with bones	small			x	250	=
18. Spinach, cooked	½ cup			x	129	=
19. Tofu, regular processed	1/3 cup			x	150	=
20. Enriched products‡				x		=
					Total: (Add lines 1-20)	

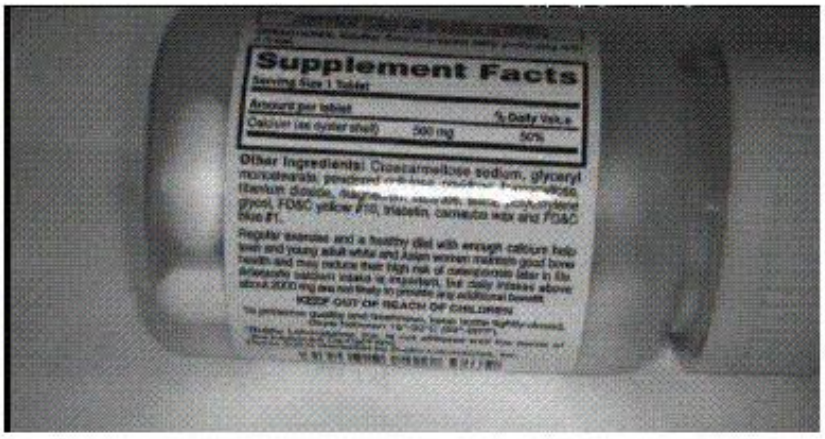
† Standard servings per week x calcium per standard serving = calcium per week.
‡ Enriched products (e.g., orange juice); provide calcium per standard serving (mg).

Average Daily Dietary Calcium (mg/day) = Total amount of Calcium (mg) Per Week ÷ 7 =	_____7
↓	
Average Daily Dietary Calcium (mg/day)	_____ mg

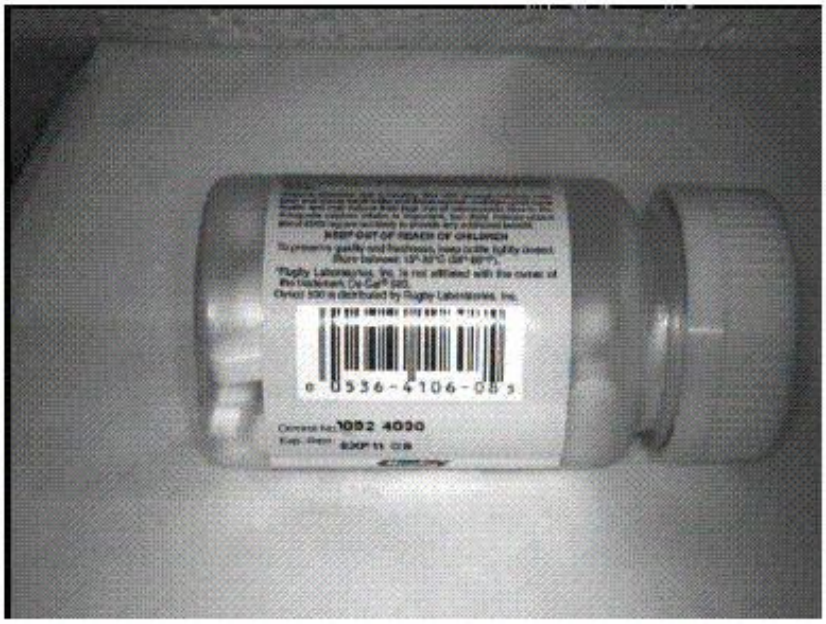
6.6 CALCIUM CARBONATE LABEL

Product: MK-0822 12
Protocol/Amendment No.: 018-04





Product: MK-0822 13
Protocol/Amendment No.: 018-04



6.7 ADDITIONAL DETAILS ON THE STATISTICAL ANALYSIS SECTION OF THE PROTOCOL

This appendix contains additional details to the data analysis section of the protocol, these details were included in the separate SAP and are now included here to eliminate the need for a separate document:

Study Participant Characteristics

Demographic and baseline characteristics of all randomized patients who received at least one dose of study medication, will be summarized for each of the treatment groups. No statistical testing will be performed to compare the treatment groups.

Summary statistics, consisting of the number of observations, mean, median, standard deviation, and range will be provided for all treated patients and by treatment group, for the following continuous variables:

- age,
- years since last menses,
- weight, height, body mass index, blood pressure and pulse rate,
- lumbar spine, total hip, femoral neck, trochanter, total body, and distal forearm BMD, by machine type,
- lumbar spine, total hip, femoral neck, trochanter, total body, and 1/3 distal forearm BMD t-scores,
- biochemical markers of bone resorption: u-NTx, s-CTx,
- biochemical markers of bone formation: s-BSAP, s-P1NP,
- indices of calcium and mineral homeostasis: s-PTH, 25-hydroxyvitamin D,
- total average daily elemental calcium intake from food, beverage and supplements.

Frequency tables for the complete set of treated patients and by treatment groups will be provided for the following categorical variables:

- stratum (no prior vertebral fracture, ≥1 prior vertebral fracture),
- age (<70 years, ≥70 years),
- race and ethnicity,
- family history of osteoporosis (yes, no),
- tobacco use (no, yes and current user/ex-user),
- alcohol use (≤7, >7 drinks per week),
- fracture history (yes, no),

- fracture history type (spine, hip, wrist or other),
- fracture history since last menses (yes, no),
- fracture history type since last menses (vertebral, hip, or non-vertebral),
- osteoporosis therapy candidacy (candidate for osteoporosis therapy, yes/no).

Two height measurements will be taken and recorded on the electronic Case Report Form (eCRF). If the 2 measurements differ by 4 mm or more, a third and fourth measurement will be obtained and recorded on the eCRF. The mean of the last 2 measurements will be used as the estimate of stature.

The number and percentage of patients with secondary diagnoses will be summarized by treatment group for each system organ class and preferred term, for terms with an incidence of at least 1% in one of the treatment groups. Similar summaries will be provided for all specific secondary diagnoses and for the number of patients with prior medications and for the number of patients with concomitant medications. Special attention will be paid to prior osteoporotic medications.

Study participants accounting will be examined by tabulation of the number of patients screened but not randomized (with the reason for not being randomized), the number of patients randomized, the number of patients completing the study, and the number of patients discontinuing the study early with the reasons for early discontinuation and the number of patients who discontinued treatment early with the reason for early termination. Since this is a pure Intention to Treat (ITT) study, patients are to be followed off-drug for the remainder of the study if they discontinue treatment early.

Definition of Compliance Measure

For each patient, two compliance measures will be calculated based on the prime therapy records provided in the eCRFs. The first one describes the compliance to treatment up to treatment discontinuation and the second during the complete study.

Compliance During the Treatment Period:

The compliance during the treatment period, until permanent discontinuation from treatment, will be defined as:

$$\text{compliance during treatment period (\%)} = 100 * \frac{\text{actual number of days with double - blind treatment intake}}{\text{expected number of days with intake in treatment period}}$$

The ”actual number of days with double-blind treatment intake” refers to the number of days on which double-blind treatment was taken.

The ”expected number of days with intake in treatment period” refers to the number of days during the treatment period on which a weekly tablet was to be taken, calculated as

the number of days from the study treatment start day to the date of the last treatment intake before the cut-off date for analysis divided by 7 (rounded upwards).

Compliance During the Complete Study:

In addition, the compliance during the complete study, including the period after permanent treatment discontinuation for patients who continue to be followed, will also be summarized and is defined as:

$$\text{compliance during the complete study (\%)} = 100 * \frac{\text{actual number of days with double - blind treatment intake}}{\text{expected number of days with intake in complete study}}$$

The "actual number of days with double-blind treatment intake" refers to the number of days on which double-blind treatment was taken.

The "expected number of days with intake in complete study" refers to the number of days during the study on which a weekly tablet was to be taken, calculated as the number of days from the study treatment start day to the date of the last contact before the cut-off date for analysis divided by 7 (rounded upwards).

Summary statistics including number of observations, mean, median, standard deviation and range, will be provided for the 2 treatment groups for each compliance measurement.

Extent of Exposure to Drug

The extent of exposure (duration and cumulative dose) to the treatment will be summarized for the treatment groups. The total duration of therapy will be defined as the period between the first day with a double-blind treatment intake and last day with a treatment intake + 7 days. Summary statistics, consisting of the number of observations, mean, median, standard deviation, and range will be provided for the 2 treatment groups.

The average total daily dose of calcium from food, beverages and supplements at each time point will also be summarized similarly.

In order to investigate the potential bias introduced by the time windows, defined below, a summary (mean, median, standard deviation [SD]) will be provided of the number of days since the start of double-blind study medication of the measurement considered in the analysis for each of the time windows for the morphometric vertebral fractures and BMD endpoints.

Additional Details on Safety Analyses:

Study Patient Population for Safety Analyses:

The "All Patients-as-Treated" (APaT) approach will be used, which includes all patients who were randomized to double-blind treatment and received at least one dose of study medication. For this analysis patients will be counted in the treatment group for the treatment they actually received: if patients received one dose of the medication they were randomized to, they will be considered in their randomization treatment group, only if they received none of their actual randomization treatment and received the other treatment they will be considered in the other treatment group.

For vital signs and laboratory safety endpoints the APaT approach will be used, including all patients who took at least one dose of double-blind study medication and had a baseline and on-treatment measurement. Missing data will not be imputed.

Clinical Safety

Summary statistics will be provided for the change from baseline at each on-treatment time point in weight, blood pressure and pulse rate, by treatment group. Height will be handled as an efficacy endpoint.

Laboratory Safety

The primary focus of laboratory data will be based, for each laboratory test, on the proportion of patients exceeding predefined limits of change from baseline and/or exceeding predefined values, at least once during treatment period, until 14 days after the last dose of double-blind treatment.

The predefined limits of change from baseline and predefined values are displayed in Table 1.

For those predefined limits characterizing a large increase from baseline and/or a large on-treatment value, a patient exceeding the predefined limit will be classified as "above limit." For those predefined limits characterizing a large decrease from baseline and/or a low on-treatment value, a patient exceeding the predefined limit will be classified "below limit." The proportion of patients exceeding the predefined limits will be presented for the treatment groups, together with the 95% confidence intervals on the difference between the treatments.

In addition to the analysis of predefined limits of change, the change from baseline at each of the on-treatment time points in laboratory tests will be summarized for each of the treatment groups. No tests will be performed on the laboratory data.

A listing of all laboratory tests performed is in the appendix.

Table 1

Predefined Limits of Change Definitions for Laboratory Tests

Laboratory Test	Predefined Limit of Change
Hemoglobin	Decrease ≥10% and <LLN Increase ≥10% and >ULN
Hematocrit	Decrease ≥10% and <LLN Increase ≥10% and >ULN
WBC count	Decrease ≥20% and <LLN Increase ≥20% and >ULN Value <LLN
Estimated neutrophils count	Decrease ≥20% and <LLN Increase ≥50% and >ULN
Estimated lymphocyte count	Decrease ≥20% and <LLN Decrease ≥20% and <1.0 $x10^3$/microL Increase ≥50% and >ULN Absolute lymphocyte count <1.0 $x10^3$/microL
Estimated eosinophil count	Increase ≥50% and >ULN Value ≥0.65 ths/mm^3
Platelet count	Decrease ≥20% and <LLN Increase ≥30% and >ULN
BUN	Increase ≥50% Increase ≥20% and BLN >ULN
Serum creatinine AST	Increase ≥50% Increase ≥20% and BLN >ULN Increase ≥50% and >ULN
ALT	Increase ≥50% and >ULN
Serum alkaline phosphatase	Increase ≥30% and >ULN
Serum sodium	Decrease ≥5 mEq/L and <LLN Increase ≥5 mEq/L and >ULN
Serum potassium	Decrease ≥0.5 mEq/L and <LLN Increase ≥0.5 mEq/L and >ULN
Serum phosphate Serum albumin	Decrease ≥30% and <LLN Increase ≥30% and >ULN Value ≤2.0 mg/dL also using value ≤1.5 mg/dL <3.0 g/dL
Serum direct bilirubin	Increase >50% and >ULN
Serum total bilirubin	Increase >50% and >ULN
Serum magnesium	Decrease >0.2 mEq/L and <LLN Increase >0.2 mEq/L and >ULN
Serum calcium	Value <8.5 mg/dL Value >10.5 mg/dL also using Value <8.0 mg/dL

Predefined Limits of Change Definitions for Laboratory Tests (Cont.)

Laboratory Test	Predefined Limit of Change
Serum corrected calcium	Value <8.5 mg/dL Value >10.5 mg/dL also using Value <8.0 mg/dL
ULN = Upper Limit of Normal. LLN = Lower Limit of Normal. BLN = Baseline (pretreatment) value.	

Analyses and Populations of the Interim Analysis

Specific safety tables provided for the safety reviews before the first formal interim analysis (70% hip fracture events) will be discussed by the DMC during the kick-off meeting and detailed in the DMC charter and will be similarly to those mentioned in 3.5.5.2.

For the 2 formal efficacy interim analyses (70% and 85% hip fracture events), analyses will be similar as those mentioned in Section 3.5.5.1 and 3.5.5.2. Analyses will focus mainly on fractures and key secondary endpoints and populations of analyses will be similar to those mentioned in Section 3.5.5.1. Additional secondary analyses as mentioned in Section 3.5.5.2 may be performed and will be determined by the DMC in their charter before unblinding. Similarly as mentioned in Section 3.5.5.1, for each interim analysis, a calendar date (cut-off) will be determined when the pre-specified number of hip fracture events are seen. Patients who did not have a clinical fracture by the time of the cut-off date will be censored at the cut-off date, unless they discontinued the study earlier, then they will be censored at the date of the last contact with the patient before the cut-off date (either telephone contact or study visit).

Subgroup Analyses

Efficacy Subgroup Analyses

Fractures

The primary and key secondary endpoints will be summarized by stratum (prior vertebral fracture, no prior vertebral fracture) and the consistency of the treatment effect across strata will be investigated as detailed in Section 3.5.5.2 by summary statistics per stratum. In addition, for the primary endpoints (morphometric vertebral, non-vertebral and hip fractures), differential treatment effects will be explored across various subgroups, e.g., age (<70 years, ≥70 years), age (<70 years, 70 to <80, ≥80 years), race (white, black, Asian, multi-racial, other), baseline BMD t-score tertiles, baseline biochemical marker tertiles, geographic region (e.g., Japan, India and, in addition, the Far East Region[1]), prior

[1] Far East Region includes China, Hong-Kong, Korea, Taiwan and Japan. If countries/sites are added to the trial, this definition may be revisited.

vertebral fracture (0, 1, >1), osteoporosis therapy candidacy (candidate for osteoporosis therapy, yes/no), participation in the lead or main cohort, and use of intranasal calcitonin (yes/no). Fracture incidence rates will be provided for each of the subgroups (by treatment) by life-table estimates for morphometric vertebral fractures and Kaplan-Meier estimates for clinical fractures. No statistical testing will be performed to compare the treatment effect across subgroups.

A patient is considered not an appropriate candidate for osteoporosis therapy treatment if she had contraindications, demonstrated tolerability issues, or had a physician concern against or was unwilling to take other osteoporosis therapy.

Bone Mineral Density

Percent change from baseline in BMD endpoints will also be summarized by osteoporosis therapy candidacy (candidate for osteoporosis therapy, yes/no) and treatment, for each time point as well as by bisphosphonate intolerance and participation in the lead or main cohort.

Percent change from baseline in BMD endpoints will also be summarized by renal function, assessed by the estimated creatinine clearance (15 to <30, 30 to <60, 60 to <90, ≥90 mL/min). The formula for the calculation of the creatinine clearance is in the exclusion criteria section.

In addition, the same subgroup analysis will be performed using the IDMS-traceable MDRD GFR calculation:

$$\text{GFR (mL/min/1.73 m}^2\text{)} = 175 \times \text{Screat}^{-1.154} \times \text{Age}^{-0.203} \ (\times 0.742 \text{ if female})$$

Similar subgroup analyses for BMD endpoints will also be performed based on 25-hydroxyvitamin D levels (<20 , ≥20 ng/mL) at baseline and at Month 12.

Safety Subgroup Analyses

Frequency tables for the number of patients with any adverse experience, drug-related adverse experiences, serious adverse experiences and who discontinued treatment due to an adverse experience will be provided by age (<70 years, ≥70 years), race (white, black, Asian, multi-racial, other), bisphosphonate intolerance, by osteoporosis therapy candidacy (yes/no), and use of azole antifungals (strong CYP3A4 inhibitors), for clinical adverse experiences, laboratory adverse experiences, skin disorders, dental disorders and investigator-reported fractures. Similar tables will also be provided specifically for Japanese patients.

Adhoc Analyses

Composite Endpoints of Fractures and Excessive Bone Loss:

As an exploratory analysis, the effect of MK-0822 on the time to first morphometric vertebral fracture or excessive bone loss (whichever comes first) will be investigated similarly as morphometric vertebral fractures. In this analysis, the time to event will be the minimum of the time to the first morphometric vertebral fracture and the date of the first Dual-energy x-ray absorptiometry (DEXA) scan on which the patient had excessive bone loss, as defined in Section 3.4.1.1 (a loss at the lumbar spine or total hip of 7% or greater compared to baseline at any point in the trial).

Furthermore, the composite endpoint of time to first hip fracture and excessive bone loss will also be investigated.

These analyses will be performed to account for patients who experienced excessive bone loss and discontinued study medication and switched to alendronate (or other osteoporosis) therapy, as instructed per protocol (Section 3.4.1.1).

Blinding/Unblinding

This study is double-blind with in-house blinding rules (i.e., blinding for patients, investigators, and Merck personnel). Throughout the entire study, the SPONSOR, investigator, patient, Central Laboratory and BMD Quality Assurance Center will remain blinded to individual treatment allocation. Individual laboratory and other results, which may identify treatment (e.g., BMD, biomarkers), will not be revealed until the end of the complete study (when all hip fracture events are observed, close-out visits have occurred and database is locked for the final analysis). A limited number of people involved in the DMC will be unblinded to treatment allocations and individual patient data.

If there are abnormalities in the blinded BMD results (e.g. excessive bone loss as detailed in the protocol), the QC Center will promptly notify both the investigator and the SPONSOR of these findings.

It is possible that there may be emergency unblinding requests for some patients due to various reasons, including adverse experiences. The timing of this type of occurrence cannot be predicted and may occur at any time during the study. Merck & Co., Inc. will make every attempt to keep the emergency unblindings to a minimum, when possible, and document these unblinded cases in the Clinical Study Report.

For the Final Analyses, all data will be screened, discrepancies resolved, and protocol violators identified before the data are frozen and unblinded to the regular clinical team. All data handling guidelines and actions will also occur prior to data unblinding according to Merck's SOP for double-blind studies with in-house blinding. At the time of unblinding, the database will be frozen in order to ensure that analyses of data in response to regulatory queries will be performed on the same data set as that which was used for the regulatory submission. The unblinded statistician for the DMC will not be involved

in any of these data review procedures. More detail about the unblinding of the interim analyses is in Section 3.5.5.6.

Technical Details on the Statistical Methods

Interval Censored Survival Methods

Morphometric vertebral fractures will be identified on x-rays taken at Month 6, Year 1 and yearly thereafter. A patient can have a morphometric vertebral fracture without sufficient symptoms or pain for it to be reported as a clinical fracture. The exact date of these fractures is therefore unknown, it is only known that the fracture occurred in the first 6 months, in the next 6 months, in the second year, etc, depending on the scan on which the fracture appeared for the first time. This data is therefore (grouped) interval-censored data with periods Month 6, Year 1, Year 2, etc. An interval-censored survival approach will be used to evaluate the treatment effect [4], [5]. A generalized linear model for binary data will therefore be used with the complementary log-log transformation of the probability of an event up to the time point, including all available data from the regularly scheduled x-rays. The model will include terms for treatment, stratum (prior/no prior vertebral fracture), and geographic region. An estimate of the hazard ratio from the model will be provided along with its 95% confidence interval.

The above model will be fit using the following SAS PROC GENMOD statement:

proc genmod data = <*data*> order = internal;

class trt period stratum region;

model cevent/n_t = trt period stratum region /d=binomial link=cloglog type3 wald;

estimate "trt" trt -1 1;

run;

For each patient multiple records are to be created for each period ("Period" taking values "0.5" for Baseline to Month 6, "1" for Month 6 to 12, "2" for Month 12 to 24, etc), until the period in which the patient has an event (first fracture per patient). The timewindows are defined below. For each record (period) an indicator variable "cevent" indicates if the patient had her first fracture in the corresponding period. "n_t" in the model is 1 for all records, "trt" is the treatment group ("Placebo", "MK-0822 50 mg"), "stratum" is the stratum (prior vertebral fracture/no prior vertebral fracture), and "region" is the geographic region.

The validity of the proportional hazards assumption will be explored (e.g. using the life-table estimates or by investigation of a similar model with treatment-time point interaction added to the model).

If a patient has 2 morphometric vertebral fractures (either in the same period or in different periods), only the first of these fractures per patient will be considered, this implies that if there is one or more fracture on the same scan the patient is considered to have a fracture event in that period, irrespective of the number of fractures seen on the scan.

If a patient has no scan at a time point, and has at least one scan on a later timepoint (with or without fracture identified), the 'missing' time point will be included as a period with no fracture identified.

Slope Analysis for Height

To assess the rate of stature loss over time a mixed model will be conducted including data of all time points (including baseline). In this analysis, the height at all time points will be analyzed by means of a mixed model (SAS proc mixed) including fixed effects for treatment, geographic region, stratum, treatment-year interaction and random intercept and slope (year). The within patient serial correlation of the values over time will be modeled by an unstructured covariance matrix. The FAS approach will be used in this analysis.

The above model will be fit using the following SAS PROC MIXED statement:

proc mixed method = reml data = <dataset> order=internal;

class an trt region yearc stratum;

model stature = trt stratum region trt*year / s;

repeated yearc / type = UN subject = an r rcorr;

random intercept year / subject = an;

Estimate 'ALN-PBO' trt*year 1 -1 / cl;

run;

Where "stature" is the stature, "an" is the allocation number, "trt" is the treatment group ("Placebo", "MK-0822 50 mg"), "year" and "yearc" are the numerical value of the year defined by the time windows in below, "stratum" is the stratum (prior/ no prior vertebral fracture), and "region" is the geographic region.

Longitudinal Models for BMD

To assess the treatment effect over time a longitudinal analysis will be conducted including data of all time points. In this analysis, the percent change from baseline in lumbar spine BMD at all during-treatment time points will be analyzed by means of a mixed model (SAS proc mixed) including fixed effects for treatment, stratum, geographic

region, and treatment-by-time interaction, as well as a random intercept for each patient. The within patient serial correlation of the values over time will be modeled by an unstructured covariance pattern. The FAS approach will be used in this analysis and no last-observation-carry-forward will be applied to the data. Estimates of the difference between MK-0822 and placebo at each time point will be provided from this model.

The above model will be fit using the following SAS PROC MIXED statement:

proc mixed method = reml data = <dataset> order=internal;

class an trt region month stratum;

model percent_change = trt stratum region trt*month / s ddfm=kenwardroger;

repeated month / type = UN subject = an r rcorr;

lsmeans trt trt*month / cl pdiff om;

run;

Where "percent_change" is the percent change from baseline in lumbar spine BMD, "an" is the allocation number, "trt" is the treatment group ("Placebo", "MK-0822 50 mg"), "month" is the month defined by the time windows below, "stratum" is the stratum (prior/ no prior vertebral fracture), and "region" is the geographic region.

The same approach will be used for the other BMD endpoints.

The weighted LS means, weighted for region and stratum size, using the OM option in SAS proc mixed, will be primarily used to summarize and plot the percent change from baseline within treatment groups at each timepoint.

The longitudinal method assumes that data are missing at random (MAR). In this study, it is expected that Missing at Random and Missing Completely at Random (MAR/MCAR) mechanisms will underlie most of the missingness, and the proportion of data missing not at random (MNAR), driven solely by unobserved values of the study endpoints, will be small. Reasons for discontinuation from the study may include lack of efficacy, clinical or laboratory adverse experiences, relocation, withdrawal of consent, protocol violations, and/or data processing issues. Missing data caused by relocation and data processing issues are likely to be MCAR. On the other hand, missing data caused by discontinuation due to lack of efficacy may belong to MAR because the discontinuation may depend on the observed efficacy outcomes. The MAR or MNAR mechanisms might each underlie the other reasons to some extent. If treatment in large part determines the loss of data for these other reasons (such as clinical or laboratory adverse experiences), the mechanism may be close to MAR because treatment assignment is an observed variable and included in the analysis model. Based on prior study results, missing data due to other reasons is relatively infrequent.

Miettenen & Nurminen Method

All tests and asymptotic confidence intervals will be based on the method found in Miettinen and Nurminen [6]. Confidence intervals will be provided to provide an estimate of the precision of the study sample estimated treatment differences in percentage of patients with at least one adverse experience. These confidence intervals are interpretable as the range of truly existing treatment differences if a large population were treated that was consistent with the sample data observed in the study.

The two-sided 95% confidence interval for a difference between 2 proportions is obtained from solving the following equation for θ:

$$\chi_{\alpha}^{2} = \frac{(\hat{p}_1 - \hat{p}_2 - \theta)^2}{\tilde{V}}$$

where

- χ_{α}^{2} is the cut point of size α from the chi-square distribution with 1 degree of freedom ($\chi_{\alpha}^{2} = 3.84$ for 95% confidence interval);
- θ is the difference between 2 population proportions (that is, $\theta = P_1 - P_2$);
- $\hat{p}_1$, $\hat{p}_2$ are the observed proportions (observed values for P_1 and P_2, respectively);
- $\tilde{V} = \left[\frac{\tilde{p}_1(1-\tilde{p}_1)}{n_1} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{n_2} \right] \frac{(n_1 + n_2)}{(n_1 + n_2 - 1)}$;
- n_1 and n_2 are the sample sizes for Group 1 and Group 2, respectively;
- $\tilde{p}_1$ = the maximum likelihood estimate of the proportion for Group 1 and is computed as $\tilde{p}_1 = \tilde{p}_2 + \theta$;
- $\tilde{p}_2$ = the maximum likelihood estimate of the proportion for Group 2 as a function of θ and under the constraint $\tilde{p}_1 - \tilde{p}_2 = \theta$.

However, since the equation does not allow for explicit solutions for θ, a numerical algorithm will be used to obtain the roots for θ. Details of the numerical algorithm can be found in Miettinen and Nurminen [6].

To account for the variable period of observation between the patients, the denominators in the above proportions (i.e. n_1 and n_2) can be replaced by the patient-years of observation. The chi-square function still has the form given above, but the variance estimate is to be calculated slightly different as explained in [6].

Definition of Geographic Region

Most statistical models will include a term for geographic region. Region will be defined as US, Europe, Asian-Pacific, and Latin-America.

Ground Rules and Data Handling Conventions

Baseline Definitions or Conventions

The baseline values for efficacy or safety endpoints used in the analysis will be the measurement closest to the target day (Day 1) for the baseline time window defined below. In case there are 2 measurements on this day, the mean of these 2 measurements will be considered in the analysis.

Time Points, Day Ranges, and Phasing of Study Periods

Since it is generally not possible for all study participants to come in for their clinical visits on the exact day specified in the protocol schedule, relative day ranges will be established for the efficacy and safety variables according to Table 3 to Table 7. The relative day of start of double-blind study medication is Day 1 (based on the prime therapy records), Day -1 is the day before and Day 2 the day after.

Table 2

Relative Day Ranges for Efficacy Analyses
Morphometric Vertebral Fractures

Time Point (Target Day)	Full-Analysis-Set and Per-Protocol Analyses
Baseline (1)	-90 to 7
Baseline to Month 6 (183)	8 to 228
Month 6 to 12 (365)	229 to 455
Month 12 to 24 (731)	456 to 821
Month 24 to 36 (1,096)	822 to 1,186
Month 36 to 48 (1,461)	1,187 to 1,551
If the study continues beyond Year 4, additional time windows will be defined similarly. All scans within a time window will be used, not only those closest to the target day.	

Table 3

Relative Day Ranges for Efficacy Analyses
BMD and Height Endpoints

Time Point (Target Day)	Full-Analysis-Set Analysis
Baseline (1)	-90 to 7
Month 6 (183)	8 to 273[†]
Month 12 (365)	274 to 548[†]
Month 24 (731)	549 to 913[†]
Month 36 (1,096)	914 to 1,278[†]
Month 48 (1,461)	1,279 to 1,643[†]
If the study continues beyond Year 4, additional time windows will be defined similarly. [†] Provided measurement was not more than 30 days after the last dose of treatment in the complete study. As a supportive analysis, an analysis including all date, also beyond 30 days after the last dose will also be performed.	

Table 4

Relative Day Ranges for Efficacy Analyses Biochemical Markers of Bone Resorption and Formation and Indices of Calcium and Mineral Homeostasis

Time Point (Target Day)	Per Protocol Analysis
Baseline (1)	-30 to 1
Month 3 (91)	49 to 133
Month 6 (183)	141 to 225
Month 9 (274)	232 to 316
Month 12 (365)	323 to 407
Month 18 (548)	506 to 590
Month 24 (731)	689 to 773
Month 30 (913)	871 to 955
Month 36 (1,096)	1,054 to 1,138
Month 42 (1278)	1,236 to 1,320
Month 48 (1,461)	1,419 to 1,503
If the study continues beyond Year 4, additional time windows will be defined similarly. Markers and selected indices are measured at baseline, Month 6, 12, 24, 36, 48, serum calcium is measured on more timepoints.	

Table 5

Relative Day Ranges for Efficacy Analyses Bone Biopsy Endpoints

Time Point (Target Day)	Full-Analysis-Set Analysis
End of Study or Month 36 (1,096)	≥518

Table 6

Relative Day Ranges for Efficacy Analyses
Health Resource Utilization and Meal Questionnaires

Time Point (Target Day)	Full-Analysis-Set Analysis	
	Health Resource Utilization Questionnaires	Meal Questionnaires[†]
Baseline (1)	-90 to 7	-90 to 1
Month 3 (91)	8 to 137	2 to 137[‡]
Month 6 (183)	138 to 228	138 to 228[‡]
Month 9 (274)	229 to 319	229 to 319[‡]
Month 12 (365)	320 to 411	-
Month 15 (457)	412 to 502	-
Month 18 (548)	503 to 593	-
Month 21 (639)	594 to 685	-
Month 24 (731)	686 to 776	-
Month 27 (822)	777 to 867	-
Month 30 (913)	868 to 958	-
Month 33 (1,004)	959 to 1,050	-
Month 36 (1,096)	1,051 to 1,141	-
Month 39 (1,187)	1,142 to 1,232	-
Month 42 (1,278)	1,233 to 1,324	-
Month 45 (1,370)	1,325 to 1,415	-
Month 48 (1,461)	1,416 to 1,506	-

If the study continues beyond Year 4, additional timewindows will be defined similarly.
[†] Meal questionnaire was only performed in the ~1,500 Lead Cohort patients.
[‡] Provided measurement was not more than 30 days after the last dose of treatment in the complete study.

Table 7

Relative Day Ranges for Safety Analyses
Clinical Safety and Laboratory Endpoints

Time Point (Target Day)	Blood Pressure, Heart Rate APaT Analysis	Weight APaT Analysis	Laboratory Tests APaT Analysis
Baseline (1)	-90 to 1	-60 to 1	-90 to 1
Month 3 (91)	2 to 137[†]	2 to 137[†]	2 to 137[†]
Month 6 (183)	138 to 228[†]	138 to 228[†]	138 to 228[†]
Month 9 (274)	229 to 319[†]	229 to 319[†]	229 to 319[†]
Month 12 (365)	320 to 456[†]	320 to 425[†]	320 to 456[†]
Month 18 (548)	457 to 639[†]	488 to 608[†]	457 to 639[†]
Month 24 (731)	640 to 822[†]	671 to 791[†]	640 to 822[†]
Month 30 (913)	823 to 1,004[†]	853 to 973[†]	823 to 1,004[†]
Month 36 (1,096)	1,005 to 1,187[†]	1,036 to 1,156[†]	1005 to 1,187[†]
Month 42 (1,278)	1,188 to 1,369[†]	1,218 to 1,338[†]	1,188 to 1,369[†]
Month 48 (1,461)	1,370 to 1,552[†]	1,401 to 1,521[†]	1,370 to 1,552[†]

APaT: All-Patients-as-Treated.
[†] Provided measurement was not more than 14 days after the last dose of treatment in the complete study.

For all endpoints, except morphometric vertebral fractures, in the event that a variable has more than one value in a day range, the value closest to the target day of the time window will be selected for the FAS analysis. If a patient has 2 values on this day, the mean of these 2 values will be taken for the analysis. For the per protocol approach, the measurement closest to the target day in the day range, which does not violate the protocol violation criteria will be selected.

For morphometric vertebral fractures, all available scans will be used to determine the presence of a fracture in a time window, not only the scan closest to the target day.

As indicated in Table 3, BMD measurements taken more than 30 days after the last dose of study medication will primarily not be included in the analyses, for results to be consistent with phase IIb and future phase III studies. For consistency with the fracture endpoints, a supportive analysis including these measurements taken long after the last dose will also be performed.

Description of Data Handling Procedures Prior to Unblinding

All study data will be screened by the data management group, the project statistician, and the clinical team before they are unblinded for the Final Analysis of the study. Identification of the patients to be excluded from the per-protocol analysis will be done

prior to unblinding of the database for the Final Analysis of the study. Protocol violations detected after unblinding of the database will be described in the study report and a justification will be provided as to why they were not detected before unblinding of the database. All data handling guidelines and procedures will also be performed prior to unblinding. The in-house unblinded database will be 'frozen' in order to insure that all analyses of data in response to regulatory queries will be performed in the same data set as was used for the regulatory submission. The unblinded statistician for the DMC interim analyses will not be involved in any of the data review procedures or determination of protocol violators.

Medical monitoring of the data will be performed in a blinded fashion by the clinical and statistical team on a regular basis during the course of the study in order to ensure the safety of the patients participating, to ensure that the safety profile of the drug is accurately represented in the clinical study data, to ensure that the efficacy data from the study are sufficient to address the primary (and key secondary) hypothesis, and to ensure that investigators complied with the protocol and with Good Clinical Practices (GCP).

Description of Protocol Violations

The per-protocol analysis set excludes patients and/or data points that represent clinically important deviations from the protocol-specified criteria, and so should not be considered a repetition of the inclusion and exclusion criteria defined in the protocol, but rather as a clinical assessment of protocol-specified deviations that may affect or confound the measures of efficacy. The per-protocol analysis will be performed as the primary approach for the biochemical marker data and may be performed as a secondary analysis for the primary endpoints (morphometric vertebral, hip and non-vertebral fractures). The per-protocol analyses for fractures will not be performed if less than 10% patients (included in the FAS approach) are protocol violators.

The listing of patients/data points who will not be included in the per-protocol analyses will be provided before frozen file.

The following rules will be applied to exclude patients/data points from the per-protocol analysis:

All Efficacy Analyses

1. Patient was <3 years postmenopause at baseline.

2. Missed >4 consecutive doses, or was <75% compliant with study medication before the biomarker measurement.

   For fractures: Data will be excluded from the PP analysis from the time point onwards when >12 consecutive doses were missed, irrespective of the timing in the study. For time point more than 6 months (180 days) from start of prime therapy data will also be disregarded from the PP analysis when the total period off-drug (not necessarily consecutive doses) is more than 25% of the time in the study. Fractures

after this time point will be disregarded and the patient will be censored at the first time point where she is considered a protocol violator[2].

3. Patient had a history of primary parathyroid disease within 2 years prior to entry (not cured by parathyroidectomy), with elevated PTH or serum calcium greater than the upper limit of normal at baseline.

4. Patient has a history of secondary hyperparathyroidism.

5. Patient has a history of hypocalcemia (<8.5 mg/dL, corrected for albumin), accompanied by an elevated PTH at baseline (>65 pg/mL).

6. Patient has a baseline TSH <0.1 microIU/mL.

7. Patient has a history of a metabolic bone disorder, other than osteoporosis (examples include Paget's Disease of Bone, osteogenesis imperfecta, osteomalacia, renal osteodystrophy).

8. Use of the following medications prior to randomization or concomitantly during the study (the doses and durations of specific treatments which will classify a patient as a protocol violator, based upon either prior or concomitant use, will be defined after data are available for review, but prior to unblinding for the first efficacy interim analysis). Patients/data points will be excluded from the start of the concomitant medication onwards.

    - anabolic steroids or glucocorticoids (≥5 mg/day prednisone or equivalent),
    - oral bisphosphonates (including alendronate, risedronate, clodronate, etidronate, ibandronate, or tiludronate),
    - I.V. bisphosphonates (including zolendronate, ibandronate, pamidronate),
    - cathepsin K inhibitor,
    - cyclosporine,
    - fluoride,
    - strontium,
    - PTH,
    - phenytoin, chemotherapy, heparin, or growth hormone,
    - estrogen ± progestin (excluding topical applications), raloxifene, lasofoxifene, or other SERM, tamoxifen, tibolone, or an aromatase inhibitor,

[2] The number of missed consecutive doses was increased from 4 to 12 because the lower number was deemed overly stringent, as it equaled only 1 month of therapy, compared to approximately 3 months for 12 doses. In addition, it was felt that overall compliance with study medication was of greater importance after the first 6 months of treatment, as the effects of odanacatib on bone mineral density and bone strength (and hence fracture risk) have been shown to be time-dependent. Lower compliance during the first 6 months of the study is less likely to have a negative impact on the overall treatment effect, as fracture rates during this time period are more likely to be influenced by the patient's baseline level of risk, and yet to be altered by treatment.

- vitamin A (retinol and/or its esters); beta-carotene use is permitted,
- vitamin D (excluding study-provided concomitant vitamin D)

Biochemical Markers and Indices of Calcium and Mineral Homeostasis

Patients will not be included in the biochemical marker analyses for one or more of the reasons listed above. In addition, for the biochemical markers of bone resorption (u-NTx, s-CTx), a value will be excluded from the analysis if the last dose of study medication prior to the value is taken >7 days prior to the value, and for biochemical markers of bone formation (BSAP, P1NP) if the last dose of study medication is taken >30 days prior to the value.

Indices of calcium and mineral homeostasis will be handled as bone formation markers (>30 days).

Datasets/Programs/Variables

The directory structure following Merck guidelines will be used for storage of the data manipulation programs, data analysis programs, output tables, graphical displays, and analysis datasets. Naming of datasets, programs and variables will be performed according to the guidelines for good programming practice.

7. ATTACHMENTS

Merck Code of Conduct for Clinical Trials

Privacy Protection of Optional Specimens for Genetic and Other Biomedical Research Collected from Clinical Trials Sponsored by Merck Sharp & Dohme Corp., a Subsidiary of Merck & Co., Inc.: A Guideline for Clinicians and Privacy Board Members

Merck*
Code of Conduct for Clinical Trials

I. Introduction

A. Purpose

Merck, through its subsidiaries, conducts clinical trials worldwide to evaluate the safety and effectiveness of our products. As such, we are committed to designing, implementing, conducting, analyzing and reporting these studies in compliance with the highest ethical and scientific standards. Protection of patient safety is the overriding concern in the design of clinical trials. In all cases, Merck clinical studies will be conducted in compliance with local and/or national regulations and in accordance with the ethical principles that have their origin in the Declaration of Helsinki.

B. Scope

Such standards shall be endorsed for all clinical interventional investigations sponsored by Merck irrespective of the party (parties) employed for their execution (e.g., contract research organizations, collaborative research efforts). This Code is not intended to apply to studies which are observational in nature, or which are retrospective. Further, this Code does not apply to investigator-initiated studies (e.g., Medical School Grant Program), which are not under the control of Merck.

II. Scientific Issues

A. Study Conduct

1. Study Design

Except for pilot or estimation studies, clinical trial protocols will be hypothesis-driven to assess safety, efficacy and/or pharmacokinetic or pharmacodynamic indices of Merck or comparator products. Alternatively, Merck may conduct outcomes research trials, studies to assess or validate various endpoint measures, or studies to determine patient preferences, etc.

The design (i.e., patient population, duration, statistical power) must be adequate to address the specific purpose of the study. Research subjects must meet protocol entry criteria to be enrolled in the study.

2. Site Selection

Merck selects investigative sites based on medical expertise, access to appropriate patients, adequacy of facilities and staff, previous performance in Merck studies, as well as budgetary considerations. Prior to study initiation, sites are evaluated by Merck personnel to assess the ability to successfully conduct the trial.

3. Site Monitoring/Scientific Integrity

Study sites are monitored to assess compliance with the study protocol and general principles of Good Clinical Practice. Merck reviews clinical data for accuracy, completeness and consistency. Data are verified versus source documentation according to standard operating procedures. Per Merck policies and procedures, if fraud, misconduct or serious GCP-non-Compliance are suspected, the issues are promptly investigated. When necessary, the clinical site will be closed, the responsible regulatory authorities and ethics review committees notified and data disclosed accordingly.

B. Publication and Authorship

To the extent scientifically appropriate, Merck seeks to publish the results of studies it conducts. Some early phase or pilot studies are intended to be hypothesis-generating rather than hypothesis testing. In such cases, publication of results may not be appropriate since the trial may be underpowered and the analyses complicated by statistical issues of multiplicity.

Merck's policy on authorship is consistent with the requirements outlined in the ICH-Good Clinical Practice guidelines. In summary, authorship should reflect significant contribution to the design and conduct of the study, performance or interpretation of the analysis, and/or writing of the manuscript. All named authors must be able to defend the study results and conclusions. Merck funding of a study will be acknowledged in publications.

III. Patient Protection

A. IRB/ERC review

All clinical trials will be reviewed and approved by an independent IRB/ERC before being initiated at each site. Significant changes or revisions to the protocol will be approved by the IRB/ERC prior to implementation, except that changes required urgently to protect patient safety and well-being may be enacted in anticipation of IRB/ERC approval. For each site, the IRB/ERC and Merck's Consent Form Review department (U.S. studies) or Clinical Research Director (non-U.S. studies) will approve the patient informed consent form.

B. Safety

The guiding principle in decision-making in clinical trials is that patient welfare is of primary importance. Potential patients will be informed of the risks and benefits of, as well as alternatives to, study participation. At a minimum, study designs will take into account the local standard of care. Patients are never denied access to appropriate medical care based on participation in a Merck clinical study.

All participation in Merck clinical trials is voluntary. Patients are enrolled only after providing informed consent for participation. Patients may withdraw from a Merck study at any time, without any influence on their access to, or receipt of, medical care that may otherwise be available to them.

C. Confidentiality

Merck is committed to safeguarding patient confidentiality, to the greatest extent possible. Unless required by law, only the investigator, sponsor (or representative) and/or regulatory authorities will have access to confidential medical records that might identify the research subject by name.

D. DNA Research

DNA sequence analyses, including use of archival specimens collected as part of a clinical trial, will only be performed with the specific informed consent of the subject. With IRB approval, an exception to this restriction on use of archival specimens may be possible (for instance, if specimens are de-identified and are not referable to a specific subject).

IV. Financial Considerations

A. Payments to Investigators

Clinical trials are time- and labor-intensive. It is Merck's policy to compensate investigators (or the sponsoring institution) in a fair manner for the work performed in support of Merck studies. Merck does not pay incentives to enroll patients in its trials. However, when enrollment is particularly challenging, additional payments may be made to compensate for the time spent in extra recruiting efforts.

Merck does not pay for patient referrals. However, Merck may compensate referring physicians for time spent on chart review to identify potentially eligible patients.

B. Clinical Research Funding

Informed consent forms will disclose that the trial is sponsored by Merck, and that the investigator or sponsoring institution is being paid or provided a grant for performing the study. However, the local IRB/ERC may wish to alter the wording of the disclosure statement to be consistent with financial practices at that institution. As noted above, publications resulting from Merck studies will indicate Merck as a source of funding.

C. Funding for Travel and Other Requests

Funding of travel by investigators and support staff (e.g. to scientific meetings, investigator meetings, etc.) will be consistent with local guidelines and practices including, in the U.S., those established by the American Medical Association (AMA).

V. Investigator Commitment

Investigators will be expected to review Merck's Code of Conduct as an attachment to the study protocol, and in signing the protocol, agree to support these ethical and scientific standards.

* In this document, "Merck" refers to Merck Sharp & Dohme Corp. and Schering Corporation, each of which is a subsidiary of Merck & Co., Inc. Merck is known as MSD outside of the United States and Canada. As warranted by context, Merck also includes affiliates and subsidiaries of Merck & Co., Inc."

Privacy Protection of Optional Specimens for Genetic and Other Biomedical Research Collected from Clinical Trials Sponsored by Merck Sharp & Dohme Corp., a Subsidiary of Merck & Co., Inc.: A Guideline for Clinicians and Privacy Board Members

1. Principles and Introduction

It is now well recognized that information obtained from studying and testing clinical specimens (i.e., blood, body fluids and/or tissue) may provide important indicators not only of the presence or absence of disease, but also of responses to medical treatments. The study of the relationships between such test results and drug efficacy is a critical component of the scientific research objectives for clinical development programs at Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc. (MERCK). MERCK recognizes that studying and testing clinical specimens offers unique opportunities to enhance our understanding of human disease and health and ultimately to aid in the discovery and development of novel, breakthrough medications targeted to populations with the greatest need.

MERCK also recognizes, however, that analyses of specimens derived from consenting patients, including for research purposes, must be undertaken with the utmost consideration for human dignity and privacy, as noted in the Declaration of Helsinki, US FDA Requirements (21 CFR 50.20, 50.25, and 50.27), the International Conference on Harmonization (ICH) E6 Good Clinical Practices Guideline, and the 1997 UNESCO Declaration on the Human Genome and Human Rights. This document outlines the approach of MERCK to privacy protection of optional specimens for genetic and other biomedical research.

2. Definitions

For the purposes of this document, the following terms will apply:

Biomarker: A biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease. A biomarker may be used to see how well the body responds to a treatment for a disease or condition.[1]

Genomic Biomarkers are measurable DNA or RNA characteristics that are indicators of normal biological or pathogenic processes and/or a response to therapeutic or other intervention [2]

Pharmacogenomics (PGx): the investigation of variations of DNA and RNA characteristics as related to drug response.[2] Also see http://www.i-pwg.org/cms/index.php?option=com_docman&task=cat_view&gid=81&Itemid=118 for guidance from the Industry – Pharmacogenomics Working Group (I-PWG) to Investigators, IRBs/IECs and Investigational Site Staff

Pharmacogenetics (PGt): the influence of variation in DNA sequence on drug response[2]

Patient-specific Identifiers: Generally defined as data fields alone or in combination that would reasonably allow a third party to identify who a patient is. Examples of these are: Patient/Subject names, date of birth, telephone #s.

Study Site: The local site of the investigation, where patients are actively screened, enrolled and studied as per the clinical protocol.

Coding of Specimens/Data: There are several categories of coding for clinical specimens and the data associated with them. See http://www.ich.org/LOB/media/MEDIA3383.pdf for additional detail on these categories. The standard method of coding used in clinical studies is single coding.

Central Laboratory/BioBank: The third-party entity that is responsible for accessioning, proper handling, and archiving clinical specimens.

3. Optional Specimens for Genetic and Other Biomedical Research: Data/Information associated with the specimen

Biomarkers, including genomic biomarkers, may be measured and analyzed by standard or novel methods to explore variations that may be related to the development and/or treatment of the diseases studied in clinical trials sponsored by Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc. (Merck). The research that we would like to perform on the specimens is considered to be critical to further advance the Merck's scientific understanding of the disease and drug responses. The research will further enable Merck to:

a) better understand disease and how to improve treatment of disease,
b) better understand how drugs work in individuals and different patient populations,
c) address emerging scientific questions that arise during the development of drugs and treatments, not known today, that would be very difficult to address if Merck did not collect a sample now, because it is not possible for the Merck, alone, to re-contact a patient directly to collect a tissue sample at a later date when the main study is finished,
d) discover and understand biological markers (biomarkers) that can be used to help understand therapeutic treatments and disease,
e) increase the chances that improved drugs may be commercially available by improving the risk benefit ratio of the treatment drug or similar drugs being developed to treat the disease in a patient or patient populations.

In order to realize and optimize the research that can be conducted with optional specimens, as described above, it is critical to link the patients' clinical information associated with the treatments in the protocol. In fact little or no research can be conducted without connecting the clinical study data to the specimen and it is unlikely the specimen would ever be used at all making any effort to collect it pointless to Merck and the patient. The clinical data allow specific analyses to be conducted for example a pharmacogenomic analysis might require knowing that specimens came from "men with type II diabetes between 20 and 50 years of age" or "children with asthma who did not respond to Drug X". In these instances, knowing gender, age and medical history and treatment outcomes are critical.

"Single coding is the current standard used in clinical research and offers additional safeguards to the subject's identifiers compared to the general healthcare confidentiality and privacy protection in everyday medical practice."[2] Consistent with this understanding, in clinical trials sponsored by Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., optional specimens and the data associated with them will be single-coded, providing the same level of privacy for clinical

research in general. In this model, the key that links the actual subject to their specimens and the data associated with them is maintained by study site personnel.

By exception, double-coding of the specifically-collected genetic specimen and/or related data analyses may be invoked. This option should only be used if there are local regulations **requiring** double-coding of genetic specimens and/or related data analyses. It is important to note that analyses may be double-coded, even if the specimen has not been double-coded. To request either of these options, complete the attached request form and submit it, along with written guidance supporting your local regulations, to your primary MERCK contact for the associated clinical trial.

4. Informed Consent

As per protocol procedures, patients/subjects should be presented with the consent form for Optional Specimens for Genetic and Other Biomedical Research at a designated visit. The consent should be administered in the standard manner, with special care to explain to the patient/subject that his/her privacy will be protected in the same way as it is provided for in the main study (unless double-coding is invoked, in which case there is slightly less risk of disclosure of the genetic research results). The individual administering consent should also carefully explain that the patient has the option to withdraw their specimens covered by the optional consent at a later date (See Section 6). Information pertaining to the administration and acquisition of the consent for Optional Specimens for Genetic and Other Biomedical Research will be captured in the Case Report Forms (CRFs) to assure that only appropriately-consented specimens are used for genetic and other biomedical research purposes. Any specimens for which such an informed consent cannot be verified will be destroyed.

5. Assembly of Kits, Specimen Collection and Handling

A designated Central Laboratory will be responsible for assembling and distributing specimen collection kits and labels both for the main study specimens and for the optional specimens. The Central laboratory will also provide the instructions on how to obtain, label, process and ship the specimens. Upon receipt by the Central Laboratory (or its associated biobank), the specimens will be processed and/or stored as specified in the contract and consistent with each subject's actual consent. MERCK will routinely monitor the condition and disposition of specimens at the biobank so that each specimen may be used appropriately.

If double-coding is agreed upon for the specifically-collected genetic specimens, then those specimens or their derivatives will be transferred to another container that contains a second unique code number. The key that links the single code to the second code will be kept in a secure place with limited access. All analyses and clinical data related to the specimen or its derivatives will be linked to the second or other codes and specifically not to the original single code.

6. Specimen Destruction Procedures for Withdrawal of Consent

Patients who request their specimens to be withdrawn are instructed in the consent form to contact the Investigator *in writing*. In the event that the medical records for the main study are no longer available (e.g., if the investigator is no longer required by regulatory agencies to retain the main study records), there will no longer be a link between the patient's personal information and their specimens. On this instance, the request for specimen destruction can not be processed. If medical records for the main study are available, the Investigator will contact MERCK using the supplied telephone contact (see Sponsor Contact Information section) and a form will be provided by MERCK to obtain appropriate information to complete specimen withdrawal. MERCK will identify specimens to be destroyed using an agreed upon form. After appropriate sign-off by both parties and affirmation of destruction, specimens will be retrieved from storage and incinerated or pyrolyzed such that DNA and other biomolecules are completely destroyed, i.e. rendered to a state such that the DNA is not able to be manipulated by standard molecular biological techniques (i.e. PCR). Any residual specimens or derivatives from the samples that have left the biobank and can be tracked will also be destroyed, but only after all main study testing is complete. A confirmatory letter will be sent from the biobank to MERCK and then later from MERCK to the investigator. It is the responsibility of the Investigator to inform the patient of completion of destruction. Any data that has been generated from the specimens before sample destruction will be maintained and can not be specifically deleted**.**

7. Conclusions

Merck recognizes both the tremendous potential, and the inherent responsibility that genetic and other biomedical research specimens provide to clinical studies. The procedures outlined in this document are intended to ensure that meaningful investigation of biomedical influences in disease and/or responses to therapies can be achieved while providing a high degree of privacy protection for patients in the study.

8. References

1. From National Cancer Institute: http://www.cancer.gov/dictionary/?searchTxt=biomarker
2. From International Conference on Harmonization: DEFINITIONS FOR GENOMIC BIOMARKERS, PHARMACOGENOMICS, PHARMACOGENETICS, GENOMIC DATA AND SAMPLE CODING CATEGORIES - E15; http://www.ich.org/LOB/media/MEDIA3383.pdf

Product: MK-0822 1
Protocol/Amendment No.: 018-04

8. SIGNATURES

8.1 SPONSOR'S REPRESENTATIVE

TYPED NAME	SIGNATURE	DATE

8.2 INVESTIGATOR

I agree to conduct this clinical study in accordance with the design outlined in this protocol and to abide by all provisions of this protocol; deviations from the protocol are acceptable only with a mutually agreed upon protocol amendment. I agree to conduct the study in accordance with generally accepted standards of Good Clinical Practice. I also agree to report all information or data in accordance with the protocol and, in particular, I agree to report any serious adverse experiences as defined in the SAFETY MEASUREMENTS section of this protocol. I also agree to handle all clinical supplies provided by the SPONSOR and collect and handle all clinical specimens in accordance with the protocol. I understand that information that identifies me will be used and disclosed as described in the protocol, and that such information may be transferred to countries that do not have laws protecting such information. Since the information in this protocol and the referenced Investigator's brochure is confidential, I understand that its disclosure to any third parties, other than those involved in approval, supervision, or conduct of the study is prohibited. I will ensure that the necessary precautions are taken to protect such information from loss, inadvertent disclosure, or access by third parties.

TYPED NAME	SIGNATURE	DATE