Statistical Analysis Plan

The Calgary Vitamin D Study

Bone Density Effects of High Dose Daily Vitamin D3 for Three Years

Principal Investigators:

David Hanley (Endocrinologist), McCaig Institute for Bone and Joint Health

Steven Boyd (Biomedical Engineer), McCaig Institute for Bone and Joint Health

1 Table of Contents

2 Introduction

2.1 Preface

Current vitamin D guidelines for older adults suggest serum 25-hydroxyvitamin D (25OHD) concentration should be at or above either 50 or 75 nmol/L [1,2]. Supplementation recommendations range from 400 to 1000 IU/day [1,2]. Higher levels of supplementation are proposed by some studies, suggesting the tolerable upper level limit of 4000 IU/day should be increased [3,4] because doses of vitamin D up to 10,000 IU/day are not considered toxic [1,5-7].

Randomized control trials of vitamin D supplementation have shown a positive influence on areal bone mineral density (aBMD) [8-10]; however, a recent systematic review reported very little evidence for the overall benefit of vitamin D supplementation on aBMD. This review concludes vitamin D supplements do not influence aBMD when baseline 25(OH)D levels are > 40 nmol/L, or when vitamin D is administered with calcium. There is a lack of studies with dose-response study designs [12].

There is considerable inconsistency in the evidence supporting the beneficial effect of vitamin D supplementation on bone health. This is exacerbated by differences in study design, dose amount, dose frequency, and the inclusion or exclusion of calcium supplementation. Furthermore, BMD outcome measures based on dual x-ray absorptiometry (DXA) is limited to density assessments and cannot assess bone microarchitecture and differences in bone compartments (cortical versus trabecular bone) that can be assessed with high resolution peripheral quantitative computed tomography (HR-pQCT).

Our aim is to understand the dose-dependent effect of vitamin D supplementation on bone microarchitecture in people over a three-year period, while ensuring adequate

calcium supplementation so that we can better determine the overall effect of vitamin D on bone health.

2.2 Purpose of the Analyses

The primary aims of this study are to assess, in a randomized clinical trial, whether supplementation of vitamin D_3 increases 1) volumetric bone mineral density as measured by HR-pQCT; 2) bone strength assessed by finite element analysis derived from HR-pQCT density and microarchitecture. Additionally, we will examine whether aBMD measured by DXA increases with vitamin D_3 supplementation. The secondary aims are to understand whether vitamin D_3 supplementation improves parameters of bone microarchitecture, balance, physical function and quality of life.

3 Methods

3.1 Funding and Ethical Approval

This clinical trial was designed by the lead investigators, DAH and SKB. Funding has been provided by Pure North S'Energy Foundation and funds are managed by the University of Calgary. The trial is registered with clinicaltrials.gov (NCT01900860) and has received a Health Canada Letter of No Objection to proceed. The Conjoint Health Research Ethics Board (CHREB) of the University of Calgary approved all procedures and participant consent was acquired prior to study initiation.

3.2 Study Design

This three-year randomized, double-blind clinical trial is designed to investigate the effects of daily vitamin D supplementation on bone quality, balance, physical function and quality of life. The goal was to have at least 300 people randomized in a 1:1:1 ratio to receive either 400, 4000 or 10000 IU vitamin D_3, cholecalciferol, taken orally once per day. We chose to test a daily dose of vitamin D, rather than (perhaps) more convenient intermittent higher dose preparations, because there is evidence that intermittent use of very high doses of vitamin D may be associated with increased risk of falls or fracture.[13,14]

3.3 Randomisation and Blinding

Upon meeting inclusion criteria, participants were randomized into one of the three study arms, with an equal number of men and women in each study arm. A statistician unrelated to the trial generated a randomization table, which was uploaded into the study database by the database developers. To ensure the allocation of participants into study arms was blinded to all participants and study staff, the randomization table was only visible to the database developers. The study participants and staff know the study arms as A, B and C, and do not know which group was receiving 400, 4000 or 10000 IU vitamin D_3.

4 Study Outcome Variables

4.1 Primary Outcome Variables

Table 1: Primary Outcome Variables

	Variable	Variable Name	Units	Follow-up
Finite Element Analysis	Failure Load –Radius	Failure_load_R	N	6,12,24,36
	Failure Load – Tibia	Failure_load_T	N	6,12,24,36
HR-pQCT	Total BMD - Radius	Tt_BMD_R	mg HA/ cm^3	6,12,24,36
	Total BMD- Tibia	Tt_BMD_T	mg HA/ cm^3	6,12,24,36

HR-pQCT = high resolution peripheral quantitative computed tomography, BMD = bone mineral density, TtBMD = total bone mineral density, R = radius, T = tibia

4.2 Secondary Outcome Variables

Table 2: Secondary Outcome Variables

	Variable	Variable Name	units	Follow-up
HR-pQCT Radius and Tibia	Trabecular BMD - radius	TbBMD_R	mg HA/ cm^3	6,12,24,36
	Trabecular BMD - tibia	TbBMD_T	mg HA/ cm^3	6,12,24,36
	Trabecular number - radius	TbN_R	1/mm	6,12,24,36
	Trabecular number - tibia	TbN_T	1/mm	6,12,24,36
	Cortical BMD – radius	CtBMD_R	mg HA/ cm^3	6,12,24,36
	Cortical BMD - tibia	CtBMD_T	mg HA/ cm^3	6,12,24,36
	Cortical porosity – radius	CtPo_R	%	6,12,24,36
	Cortical porosity - tibia	CtPo_T	%	6,12,24,36
DXA	Total Hip BMD	TH_aBMD	g/cm^2	12,24,36
Protocol Visit Sheet	Balance (mean of 3)	Balance_mean		12,24,36
	Timed up and go (mean of 3)	UpandGo_mean	Sec	12,24,36
	Grip strength (mean of 3)	Grip_mean	Kg	12,24,36
	Mental Health	Sf36_mcs		3,6,12,24,36
Lab	Serum 25(0H)D	Vitamind	nmol/L	3,6,12,18,24,30,24,3
	Plasma CTx	Beta_cross	ng/L	12,24,36
	Serum Parathyroid Hormone	Parathyroid_hormone	ng/L	3,6,12,18,24, 30,24,36

HR-pQCT = high resolution peripheral quantitative computed tomography, BMD = bone mineral density, 25(OH)D = 25-hydroxyvitamin D, CTx = C-telopeptide of type 1 collagen, TbBMD = trabecular bone mineral density, TbN = trabecular number, CtBMD = cortical bone mineral density, CtPo = cortical porosity, SF-36 = Short Form Health Survey questionnaire, R = radius, T = tibia, TH = total hip, aBMD = areal bone mineral density

5 Sample Size

- The sample size estimation is based on two of the four primary outcomes variables, which are Tt.BMD for the tibia and radius.
- All four primary outcome variables are highly correlated, and each variable will be tested at the alpha level of 0.025, which is the traditional alpha level of 0.05 corrected for multiple comparisons given that the four outcome variables are highly correlated. A Bonferroni correction for alpha would use α/n but this assumes that the outcome variables are independent. Since these outcome variables are not we will use the correction $\alpha/\sqrt{n}$.
- The sample size is based on a one-way analysis of variance using a single p-value.
- Differences between the three groups will then be described using mean values and 95% confidence intervals calculated from this analysis.
- Data from our population based prospective cohort study of post-menopausal women, showed that there was a HR-pQCT-derived TtBMD declined between 3% (tibia) and 7% (radius) over the 5 years [15].
- Assuming that this decrease was linear over this short period of time, we can assume that we can expect at decrease of approximately 1.8% decrease over three years for the TtBMD Tibia and 4.2% for the Radius.
- These are somewhat larger than values from published data for placebo groups from a previous RCT exploring vitamin D supplementation on BMD in the hip and total body using DXA [16].
- In an RCT examining treatment with osteoporotic medication or placebo in postmenopausal women with low bone density, in which morphologic changes were

assessed using HR-pQCT at the distal radius and distal tibia, women taking the placebo (representing normal bone aging) had an annual total bone loss of up to 2% at the radius and 0.5% in the tibia [17].

- Thus, for our sample size calculations we have allowed for a decrease in the total volumetric BMD (TtBMD, tibia and radius) in the 400 IU dose group to range from 2% to 6% over the three years of the study.
- In keeping with the primary aim of this study the sample size will be based on the ability of this study to detect a clinically relevant dose-dependent effect of vitamin D supplementation, should this exist. For the 4,000 IU group this is considered as improving the rate of decrease by 50% or more (1% to 3%) and for the 10,000 IU group this is considered to be arresting the rate of decline or even improving the values.
- Using a sub-section of our large population-based cohort [18] aged 55-70 years, the mean TtBMD at the tibia was 283 (SD 57) mg HA/cm^3. We will be testing the mean difference between the three-year value and the baseline value in each of the three groups.
- In order to estimate the standard deviation for the change score, we assumed from previous data the measurements taken one year apart would be highly correlated, so we calculated the SD of the change score using a correlation of 0.95 to estimate the covariance, which is quite conservative.
- We allowed the power to vary from 80% to 95% and the number of participants needed in each group for each scenario are presented in Table 3 below.

Table 3: Estimated Number of participants needed in each group to observed hypothesized differences at an alpha level of 0.025 with the given power

Treatment	Baseline Tt.BMD		Loss		Tt.BMD	Change		Number of Participants needed in each group for the given power			
Group	Mean	SD	Annual	3 yr	3 yrs	3 yr	SD	80%	85%	90%	95%
400 IU	283	57	2%	6%	266.0	17.0	17	25	28	31	37
4,000 IU	283	57	1%	3%	274.5	8.5	17	25	28	31	37
10,000 IU	283	57	0	0	283.0	0.0	17	25	28	31	37
400 IU	283	57	1.70%	5%	268.8	14.2	17	35	39	44	53
4,000 IU	283	57	0.83%	2.5%	275.9	7.1	17	35	39	44	53
10,000 IU	283	57	0	0	283.0	0.0	17	35	39	44	53
400 IU	283	57	1.3%	4%	271.7	11.3	17	54	60	69	82
4,000 IU	283	57	0.7%	2%	277.3	5.7	17	54	60	69	82
10,000 IU	283	57	0	0	283.0	0.0	17	54	60	69	82
400 IU	283	57	1.2%	3.3%	274.5	10.2	17	66	73	84	100
4,000 IU	283	57	0.6%	1.8%	278.7	5.1	17	66	73	84	100
10,000 IU	283	57	0	0	283.0	0.0	17	66	73	84	100
400 IU	283	57	1.1%	3.3%	274.5	9.3	17	79	88	100	120
4,000 IU	283	57	0.6%	1.8%	278.7	4.7	17	79	88	100	120
10,000 IU	283	57	0	0	283.0	0.0	17	79	88	100	120
400 IU	283	57	1.0%	3%	274.5	8.5	17	94	105	120	144
4,000 IU	283	57	0.5%	1.5%	278.7	4.2	17	94	105	120	144
10,000 IU	283	57	0	0	283.0	0.0	17	94	105	120	144
400 IU	283	57	0.7%	2%	277.3	5.7	17	206	231	264	317
4000 IU	283	57	0.3%	1%	280.1	2.8	17	206	231	264	317
10,000 IU	283	57	0	0	283.0	0.0	17	206	231	264	317

- Sample size calculations, based on these assumptions showed that for an annual loss of more than 1.0% in the 400 IU group, there will be sufficient power to detect the hypothesized differences should these exist. For example, when the annual loss in the 400 IU group is 1.1% then 79 participants in each group (237 in total) will be required to achieve 80% power at the alpha level of 0.025, whereas if the annual loss in this group is 1.0% then 94 participants in each group (282 in total) would be required.
- Therefore, we plan on recruiting 84 patients per group (total 252), which would give 90% power at an alpha level of 0.025 if the annual loss in the 400 IU group is 1.2%, and allowing for 20% attrition this will require recruiting 100 participants in each group for a total of 300.
- In our laboratory, we have established the reproducibility of HR-pQCT parameters, reporting total bone density (TtBMD) reproducibility of 0.6% [19]. This is in part due to the 3D image registration techniques that have been developed to maximize our ability to detect change [19]. By implementing 3D image registration, we expect the three-year changes observed in this study to be larger than scanner precision and highly reproducible, allowing us to detect changes as small as 1% in TtBMD.

6 General Considerations

6.1 Timing of Analyses

- The final analysis will be performed after the final patient enrolled has completed follow-up, the database is completed, cleaned and locked.

6.2 Analysis Populations

6.2.1 Efficacy Analysis Population

- Modified Intent to Treat: All subjects who received any study drug and who participated in at least one post-baseline assessment.

6.2.2 Safety Population

- All subjects who were randomised and therefore received any study drug.

6.3 Covariates and Subgroups

- No subgroup analysis will be done for the primary efficacy analysis.
- No other covariates will be included in the primary efficacy analysis.
- Future analysis will examine factors such as the effect of treatment for varying baseline level of serum vitamin D. These analyses will be exploratory, since the sample size was not powered to detect any additional interactions.
- Other exploratory analyses will be planned after the conclusion of the primary efficacy and safety analyses ae completed.

6.4 Missing Data

- From pilot study data, it is not anticipated that there will be many missing data points in the primary and secondary outcome variables.
- In descriptive statistics missing data will be quantified per variable (%).
- Potential patterns of missing data will be examined.
- Missing data will be taken into account using the linear random effects models which is considered better than using the Last Observation Carried Forward (LOCF) approach [20].

6.5 Interim Analyses and Data Monitoring

- There will be no interim analysis.

6.6 Multi-centre Studies

- This is a single centre trial.

6.7 Multiple Testing

- The four primary outcome variables likely to be (highly) correlated, Bonferroni correction (α/k) where k is the number of tests applies to independent tests. Tukey suggested using $\alpha/\sqrt{k}$ when outcome variables are correlated but the correlation is unknown. Therefore, we could treat all four primary outcomes as equally important and test them at α = 0.025.
- Rather than solely relying on p-values for the interpretation of the results. The results will be presented using predicted means with 95% CI of the fixed effects (time point and treatment) from each random effects regression model.
- All the secondary outcomes will be tested with α = 0.05. However, keeping in mind the large number of secondary outcomes examined, the large probability of

spuriously significant results at the 5% level of significance (and even the 1% level of significance) will be kept in mind.

7 Descriptive Analysis

- Initial statistical analyses will describe baseline demographic, medical, and lifestyle characteristics of the study participants by randomized group at baseline using the mean and standard deviation for approximately normally distributed variables and n (%) for categorical variables. The intent to treat population will be used.
- The four primary outcome variables will be described using the mean, standard deviate and % missing values by treatment group (intent to treat) at each time point (baseline, 6 months, 12months, 24 months and 36 months).
- The secondary outcome variables will be described using the mean, standard deviate and % missing values by treatment group at each point in time (baseline, 6 months, 12months, 24 months and 36 months).
- The Lab Safety value variables will be described using the mean, standard deviate and % missing values by treatment group at each point in time (baseline, (3 months) 6 months, 12months, (18 months) 24 months (30 months) and 36 months). 25-hydroxyvitamin D, PTH = Parathyroid hormone and CTx = C-telopeptide of type 1 collagen. Laboratory normal range is: 25(OH)D: 80 – 200 nmol/L; PTH: 7 – 37 ng/L; CTX: 0 – 400 ng/L.

8 Efficacy Analyses

- A constrained linear mixed effects model [21] will be used to analyse each outcome (both primary and secondary) variable. Baseline values will be constrained to be the same in each of the three treatment groups, since these were measured prior to randomization. The dependent variable will include all measurements taken post baseline.
- Visual inspection of individual profile plots will be done to assess the potential of a non-linear effect and if evident will be included in the model as a quadratic treatment effect.
- Fixed effects will include, time, treatment group and a potential treatment by time interaction (which if significant will be indicative of efficacy).
- Potential Random effects will include both intercept and slope. The correlation within subjects will be modelled as an autoregressive error of order 1.
- Diagnostic residual plots will be examined for deviation from the assumptions underlying the model and if necessary an appropriate transformation (such as a logarithm) of any deviant outcome variable will be applied. If such a transformation is necessary, the results will be presented in the original units.
- We will start with the full model. Unnecessary quadratic terms will be removed as will unnecessary random slopes or intercept terms to arrive at a parsimonious well-fitting model for each outcome variable.
- Non-significant treatment by time interactions will be retained in the model and reported to describe the lack of observed treatment effect.

- Missing values will be accounted for by using the linear mixed effects model which is considered a superior method to the of using LOCF [20]
- The results of the regression modelling will be described in table format where the Likelihood Ratio Statistic (LRS) for the quadratic term on 1 df (before removal in cases where this term was unnecessary) and the LRS on 2 df for the treatment group by time interaction, which yield the p-value for the treatment effect.
- Since coefficients for both quadratic terms and interaction terms are hard to interpret, results of the regression models will also be presented as the mean (with 95% CI) fixed effects calculated from the coefficients.

9 Safety Analyses

9.1 Population

- All randomized patients who receive at least one dose of the study drug.

9.2 Pre-specified Safety Outcomes

- The pre-specified safety outcomes are divided into three groups: Biochemical parameters, occurrence of AEs (deaths, serious AEs) and AE of special interest (nephrolithiasis, cancer, falls and fractures). The incidence of infections and upper respiratory tract infections were exploratory outcomes.

300 Table 4: Pre-specified Safety Outcomes

Biochemical Parameters	
	Hypercalcemia
	Hypercalciuria
	Creatinine >133 umol/L
	estimated Glomerular Filtration Rate (eGFR) decline of >10 mL/min
	Aspartate aminotransferase (AST) or Alanine aminotransferase (ALT) > 1.5x ULN[a]
Clinical Adverse Events (AE)[b]	
All Clinical AEs	Neurologic
Serious AEs	Ophthalmologic
	Otalaryngologic
	Cardiovascular
	Pulmonary
	Gastrointestinal
	Genitourinary
	Endocrine
	Hematologic
	Dermatologic
	Musculoskeletal
	Psychiatric
	Other [c]
AEs of Special Interest	
	Falls
	Low-trauma fractures
	Nephrolithiasis
	Non-skin cancer [d]
	Skin cancer
	Infections
	Upper respiratory tract infections

ULN = upper limit of normal,

[a] AST ULN = 32 IU/L for females and 40 IU/L for males, ALT ULN = 40 IU/L for females and 60 IU/L for males,

[b] AEs and serious AEs defined using the standard International Conference on Harmonization Good Clinical Practice definition

[c] AEs that do not localize to a single organ system (e.g. diffuse infectious symptoms, generalized allergic reactions, electrolyte abnormalities, fatigue, insomnia, weight changes)

[d] includes melanoma

301

9.3 Statistical Analysis of Safety Parameters

9.3.1 Baseline Descriptive Statistics

- Initial statistical analyses will describe baseline demographic, medical, and lifestyle characteristics of the study participants by randomized group at baseline using the mean and standard deviation for approximately normally distributed variables and n (%) for categorical variables.

9.3.2 Graphical Analysis

- The distributions of the continuous variables will be illustrated using boxplots. Box plots of three-year changes in serum 25-hydroxyvitamin D, serum calcium, serum creatinine, and 24-hour urine calcium in healthy adults taking vitamin D 400IU, 4000 IU, or 10000 IU/day. Boxes show medians and interquartile ranges. The whiskers show the adjacent values, which indicate where approximately 99% of the values of the data lie. Horizontal dashed lines represent the upper limit of the normal range for serum calcium, 133 µmol/L for serum creatinine, and 24-hour urine calcium excretion of 7.5 mmol/day.

9.3.3 Proportion of Participants with each AE

- For each AE, the total number of occurrences in each treatment arm will be tabulated. The proportion of participants in each treatment group who experienced each AE will determined and examined formally (for pre-specified safety outcomes, provided the overall prevalence of the AE greater than or equal to 4% and less than or equal to 96%) for between treatment group differences for trend in proportions using logistic regression.

9.3.4 Incidence Rate Differences

- The proportion of healthy individuals experiencing relevant biochemical and clinical adverse events (AEs) while taking vitamin D 400, 4000, or 10000 IU/day for three years, using 400 IU/day as the referent**.** Incidence rates reflect the number of participants experiencing the event per person-year of follow-up. Error bars will represent 95% confidence intervals. *When calculating the incidence of adverse events each subject will only be counted once and any repetitions will be ignored; the denominator will be the total population size.*
- P-values <0.05 will be considered statistically significant and were not adjusted.

10 Reporting Conventions

- P-values ≥0.001 will be reported to 3 decimal places; p-values less than 0.001 will be reported as "<0.001".

- The mean, standard deviation, and any other statistics other than quantiles, will be reported to one decimal place greater than the original data. Quantiles, such as median, or minimum and maximum will use the same number of decimal places as the original data. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 3 significant figures.

11 Software

- Statistical analyses will be conducted using the R project for Statistical Computing (R Studio, version 1.0.143). R Markdown will be used to produce reproducible statistics documentation including tables and graphs.

12 References

[1] IOM 2011 Dietary reference intakes for calcium and vitamin D, The National Academies Press, Washington, DC.

[2] M.F. Holick, N.C. Binkley, H.A. Bischoff-Ferrari, C.M. Gordon, D.A. Hanley, R.P. Heaney, et al., Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline, J. Clin. Endocrinol. Metab. 96 (2011) 1911–1930. doi:10.1210/jc.2011-0385.

[3] R.P. Heaney, K.M. Davies, T.C. Chen, M.F. Holick, M.J. Barger-Lux, Serum 25 hydroxy cholecalciferol response to oral dosing with cholecalciferol vitamin D doses response relationships, 77 (2003) 204–210.

[4] R. Vieth, P.C. Chan, G.D. MacFarlane, Efficacy and safety of vitamin D3 intake exceeding the lowest observed adverse effect level, Am J Clin Nutr. 73 (2001) 288–294. doi:10.3945/ajcn.116.136697.

[5] J.N. Hathcock, A. Shao, R. Vieth, R. Heaney, Risk assessment for vitamin D, Am J Clin Nutr. 85 (2007) 6–18. doi:10.1097/01.mog.0000196150.36701.c2.

[6] R. Rizzoli, S. Boonen, M.L. Brandi, O. Bruyère, C. Cooper, J.A. Kanis, et al., Vitamin D supplementation in elderly or postmenopausal women: a 2013 update of the 2008 recommendations from the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO), in: Curr Med Res Opin, 2013: pp. 305–313. doi:10.1185/03007995.2013.766162.

[7] A.C. Ross, J.E. Manson, S.A. Abrams, J.F. Aloia, P.M. Brannon, S.K. Clinton, et al., The 2011 Report on Dietary Reference Intakes for Calcium and Vitamin D from the Institute of Medicine: What Clinicians Need to Know, J. Clin. Endocrinol. Metab. 96 (2011) 53–58. doi:10.1210/jc.2010-2704.

[8] M.E. Ooms, J.C. Roos, P.D. Bezemer, W.J. van der Vijgh, L.M. Bouter, P. Lips, Prevention of bone loss by vitamin D supplementation in elderly women: a randomized double-blind trial, J. Clin. Endocrinol. Metab. 80 (1995) 1052–1058. doi:10.1210/jcem.80.4.7714065.

[9] I.R. Reid, A.M. Horne, B. Mihov, G.D. Gamble, F. Al Abuwsi, M. Singh, et al., Effect of monthly high-dose vitamin D on bone density in community-dwelling older adults substudy of a randomized controlled trial, Journal of Internal Medicine. 383 (2017) 146. doi:10.1111/joim.12651.

[10] H.M. Macdonald, A.D. Wood, L.S. Aucott, A.J. Black, W.D. Fraser, A. Mavroeidi, et al., Hip bone loss is attenuated with 1000 IU but not 400 IU daily vitamin D3: A 1-year double-blind RCT in postmenopausal women, J. Bone Miner. Res. 28 (2013) 2202–2213. doi:10.1002/jbmr.1959.

[11] I.R. Reid, M.J. Bolland, A. Grey, Effects of vitamin D supplements on bone mineral density: a systematic review and meta-analysis, Lancet. 383 (2014) 146–155. doi:10.1016/S0140-6736(13)61647-5.

[12] L.M. Smith, J.C. Gallagher, Dietary vitamin D intake for the elderly population: update on the recommended dietary allowance for vitamin D, Endocrinol. Metab. Clin. North Am. 46 (2017) 871–884, http://dx.doi.org/10.1016/j.ecl.2017.07.003.

[13] K.M. Sanders, A.L. Stuart, E.J. Williamson, J.A. Simpson, M.A. Kotowicz, D. Young, et al., Annual high-dose oral vitamin D and falls and fractures in older women: a randomized controlled trial, JAMA. 303 (2010) 1815–1822. doi:10.1001/jama.2010.594.

[14] H. Smith, F. Anderson, H. Raphael, P. Maslin, S. Crozier, C. Cooper, Effect of annual intramuscular vitamin D on fracture risk in elderly men and women--a population-

382 based, randomized, double-blind, placebo-controlled trial, Rheumatology (Oxford).
383 46 (2007) 1852–1857. doi:10.1093/rheumatology/kem240.
384 [15] L.A. Burt, A.L. Menard, H.M. Macdonald, D.A. Hanley, S.K. Boyd, A five-year
385 longitudinal study of site-specific changes in bone quality in the calgary population
386 based cohort: an HR-pQCT study, J. Bone Miner. Res., 28 (Suppl. 1). Available at:
387 http://www.asbmr.org/asbmr-2013-abstract-detail?aid=bf061225-0796-4f92-99c9-
388 51edbbbfc58f.
389 [16] Dawson-Hughes, S.S. Harris, E.A. Krall, G.E. Dallal, Effect of calcium and vitaminD
390 supplementation on bone density in men and women 65 years of age or older,
391 N.Engl. J. Med. 337 (1997) 670–676,
392 http://dx.doi.org/10.1056/NEJM199709043371003.
393 [17] E. Seeman, P.D. Delmas, D.A. Hanley, D. Sellmeyer, A.M. Cheung, E. Shane, et al.,
394 Microarchitectural deterioration of cortical and trabecular bone: differing effects of
395 denosumab and alendronate, J. Bone Miner. Res. 25 (2010) 1886–1894.
396 doi:10.1002/jbmr.81.
397 [18] H.M. Macdonald, K.K. Nishiyama, J. Kang, D.A. Hanley, S.K. Boyd, Age-related
398 patterns of trabecular and cortical bone loss differ between sexes and skeletal
399 sites: a population-based HR-pQCT study, J. Bone Miner. Res. 26 (2011) 50–62,
400 http://dx.doi.org/10.1002/jbmr.171.
401 [19] J.A. MacNeil, S.K. Boyd, Improved reproducibility of high-resolution peripheral
402 quantitative computed tomography for measurement of bone quality, Med Eng
403 Phys. 30 (2008) 792–799. doi:10.1016/j.medengphy.2007.11.003.1

[20] Molenberghs G, Thijs H, Jansen I, Beunckens, C. Analyzing incomplete longitudinal clinical trial data. *Biostatistics* (2004), 5, 3, *pp.* 445–464 doi: 10.1093/biostatistics/kxh001

[21] Coffman CJ, Edelman D, Woolson RF. To condition or not condition? Analysing 'change' in longitudinal randomized controlled trials. BMJ Open 2016;6:e013096. doi:10.1136/bmjopen-2016-013096

13 Explanation for the Brevity of the Statistical Analysis Plan at Trial Initiation

We started the trial with a very simple analysis plan, focused on our main goal of determining whether there were different effects of the three levels of Vitamin D dose on bone parameters, as measured by HR-pQCT. We recognized from the outset that the trial was going to generate a very large database, and that we would need expert help in finalizing an appropriate statistical analysis plan. We therefore advertised to secure a statistician as a vital team member. Unfortunately, we were not able to find such an individual at the time of trial initiation, and it was not until we were into the last half of the trial before we were able to recruit Dr. M.S. Rose to our investigator team as our statistics expert; she then helped us design what we feel is an appropriate way to analyze our data.

Following advice from our statistician co-investigator, we have limited the number of primary outcomes obtained from HR-pQCT measurements to total volumetric bone density (Tt.BMD) at the radius and tibia, and calculated failure load (Finite Element Analysis) at those sites.

14 Summary of Changes to the Statistical Analysis Plan

The statistician (M.S. Rose) who wrote this Statistical Analysis Plan joined the Team after the original protocol was approved. No changes were made to the Statistical Analysis Plan following the addition of the statistician (M.S. Rose) and our plan was finalized prior to study completion and unblinding of the data.