a) All differences from baseline statistically significant in all 3 treatment groups. Both active treatment groups statistically significant from placebo at both time points (p=0.044 at Month 12 and 0.031 at endpoint. LY333334 40 µg group BMD statistically significantly greater than 20 µg group at both time points (p=0.044 at Month 12 and 0.031 at endpoint).
b) Both LY 333334 groups greater than baseline and placebo at both time points, p<0.001 for each comparison; LY333334 groups significantly different from each other at both time points, p<0.001.
c) Both LY333334 groups greater than baseline and greater than placebo at both time points, p<0.001 for each comparison. LY333334 groups differ from each other at month 12 and at end (p=0.14 and 0.38, respectively).
d) All differences from baseline significant in all 3 treatment groups, p<0.001 for each comparison. Each LY333334 treatment group BMD > placebo at each time point (p<0.001 for each comparison). Differences between the two LY333334 groups were significant at endpoint only (p=0.006).
e) Each active LY333334 group grater than baseline and greater than placebo at both time points (p<0.001 for each comparison). Differences between the two LY333334 groups significant at both time points (p<0.001 for each comparison).
f) There was no statistically significant difference between either LY333334 treatment group and placebo, or between the two LY333334 groups, at either time point. Decreases from baseline were significant within the placebo group at each time point (p=0.038 and 0.015 at month 12 and endpoint, respectively). Changes from baseline were significant only for the decrease within the 40 µg group at endpoint (p=0.035).
g) All within-group decreases from baseline in distal 1/3 radius BMD were significant at both time points (p<0.001 for each comparison). The numerically greater decreases between the LY333334 20 µg group and placebo were not statistically significant (p=0.079 at month 12 and p=0.088 at endpoint). However, the differences between the 40 µg group and placebo were statistically significant at both time points (p<0.001).
h) BMC decrease from baseline to endpoint significant in placebo (p<0.05). BMC increases from baseline significant at endpoint for both LY333334 groups (p=0.001 for the 20 µg group and p<0.001 for the 40 µ g group). At endpoint p<0.001 for each LY333334 group vs placebo.

Comments: LY333334 treatment increased BMD at all skeletal sites, except the ultradistal and distal radius. The increases were consistently greater at treatment end than at month 12, as well as in the 40 µg treatment group, compared to the 20 µg group.

The increases in total body BMC in both LY333334 groups are reassuring. The placebo group lost total body BMC by study end. Not as reassuring is the loss of BMD at the distal radius, in what appears to be a dose-dependent manner. This is of concern because of the high concentration of cortical bone at that skeletal site. Although the differences between placebo and the 20 µg LY333334 group (the indicated dose; the differences between placebo and the 40 µg group were highly significant) fell short of statistical significance, the p-values were <0.10 at each time point. The concern relates to safety, rather than efficacy, placing these p-values (if they are to be used at all) in a different perspective.

Mitigating these concerns is the fact that, thus far, there have been numerically fewer wrist fractures in the LY333334 treatment groups of GHAC than in placebo. In addition, a plausible explanation for the reduction of BMD is provided by the quantitative CT pilot study, described below. At the proximal radius, treatment with LY333334 may result in an increase in periosteal and endosteal circumference (with an increase in bone area)

without a corresponding increase in bone mineral content. This would reduce the observed BMD, while increasing computed parameters of bone strength. Further evaluation of the incidence of post-treatment wrist fractures is currently part of the sponsor's long-term post-treatment study.

At all sites, except the lumbar spine, the increases in BMD are no greater than have been generally observed following 12-24 months of treatment with alendronate, 10 mg/day. It is likely that the increased efficacy of LY333334 at the spine is due to the high concentration of trabecular bone at this site.

Subgroup analysis

The sponsor examined selected subpopulations to confirm that there is a significant effect of LY333334 on lumbar spine BMD within each subgroup; the intent of this analysis was not to search for differential treatment effects.

Similar to the subgroup analysis of vertebral fractures, the sponsor examined the effects of age, body weight, BMI, baseline spine BMD, baseline biochemical markers of bone turnover, baseline endogenous PTH levels, occurrence of study dose reduction (y/n), and occurrence of 4-6 hr post-dose hypercalcemia (y/n).

The results of the subgroup analysis are provided in detail in Table GHAC.11.37 of the NDA. A review of the data showed that treatment comparisons (LY333334 vs PBO) were statistically significant within any of the above subgroups ($p<0.03$ or less for all comparisons).

Statistically significant interactions were found in some of these subgroups. These are described in detail in the submission. Treatment effects were enhanced in the middle and highest age tertiles (relative to the lowest), in the lowest and middle tertiles of BMI (relative to the highest), and in patients with spinal BMD T-scores < -2.5 (relative to those with T-scores > -2.5; the increases were relative, not absolute). In addition, the BMD responses tended to be greater in patients with higher baseline levels of bone turnover markers, in patients with low baseline endogenous PTH levels, and in patients with high 4-6 hour post-dose serum calcium levels.

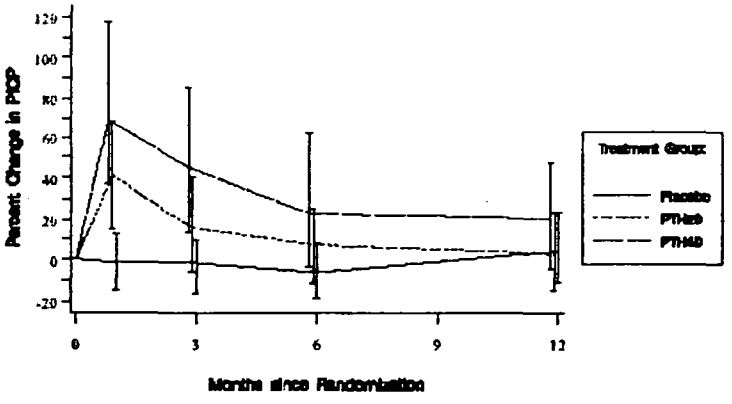
Comments: As the sponsor indicates, because of the substantial BMD treatment effects of the drug, the treatment-by-subgroup analysis had power to detect relatively weak interactions, many of which may not be clinically relevant. Certain of the results of this analysis might be used to generate hypotheses for future studies. The important result of this analysis is that, in all subgroups, treatment with LY333334 (either dose) produced statistically significantly greater increases in lumbar spine BMD than were achieved in the placebo group.

B.1.8.5.4 Biochemical markers of bone formation and resorption, and calcium regulation hormones

To characterize the action and time course of LY333334 treatment on bone remodeling, the sponsor measured levels of two markers of bone formation [serum bone-specific alkaline phosphatase (BSAP), and serum procollagen I carboxy-terminal propeptide (PICP)] and two markers of bone resorption [urinary N-telopeptide (NTX) and urinary free deoxypyridinoline]. In addition, 1,25-dihydroxyvitamin D was measured in a subset of approximately 500 patients at baseline and at Months 1, 3, 6, and 12, and at Early Discontinuation or study closeout. Demographic and other baseline characteristics of the subset of patients who had these assessments are provided in Table GHAC.14.2 of the NDA. A by-visit analysis was also performed.

Results:

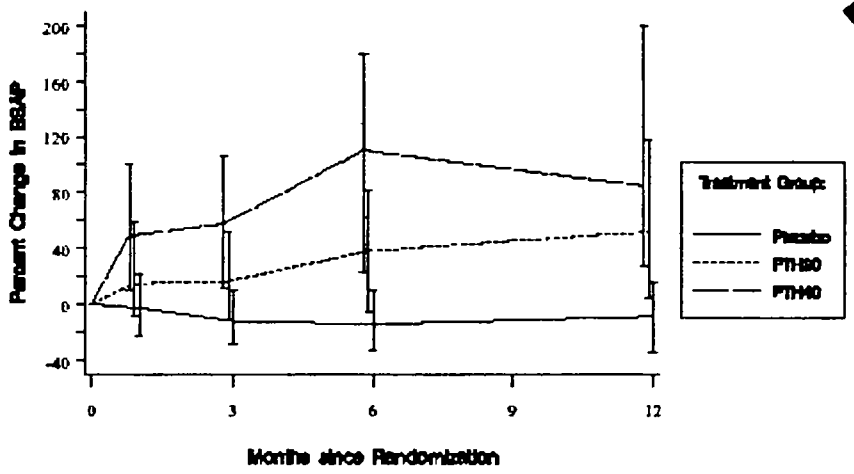
There were prompt, dose-related increases in BSAP and PICP in both LY333334 treatment groups. PICP, the earliest indicator of stimulation of osteoblast function, peaked at Month 1; and the median PICP level remained significantly higher than placebo at Months 3 and 6 in the 20 µg group and at months 1, 3, 6, and 12 in the 40 µg group. In the 20 µg group, the median PICP was slightly below baseline at Month 12. At all time points after baseline, the median % increase for the 40 µg group was statistically significantly higher than in the 20 µg group. There were no statistically significant changes in median PICP levels, compared to baseline, in the placebo group. Further analyses are provided in the NDA. The data are depicted in the following figure:



There was a small, statistically significant decrease from baseline in BSAP levels in the placebo group ($p<0.001$) at all scheduled time points except Month 3. In contrast, serum levels of BSAP increased significantly from baseline in both

LY333334 treatment groups ($p<0.001$) at all time points except for study end in the 20 µg group ($p=0.158$ at that time point; the p-value for the 40 µg group vs placebo at study end was <0.001). The increase in BSAP appeared to reach a plateau between Months 6 and 12. At all time points after baseline, both LY333334 doses produced statistically significantly greater increases in BSAP than were found in the placebo group ($p<0.001$). At all times after baseline, the increase in BSAP in the 40 µg group was greater than in the 20 µg group ($p \leq 0.015$).

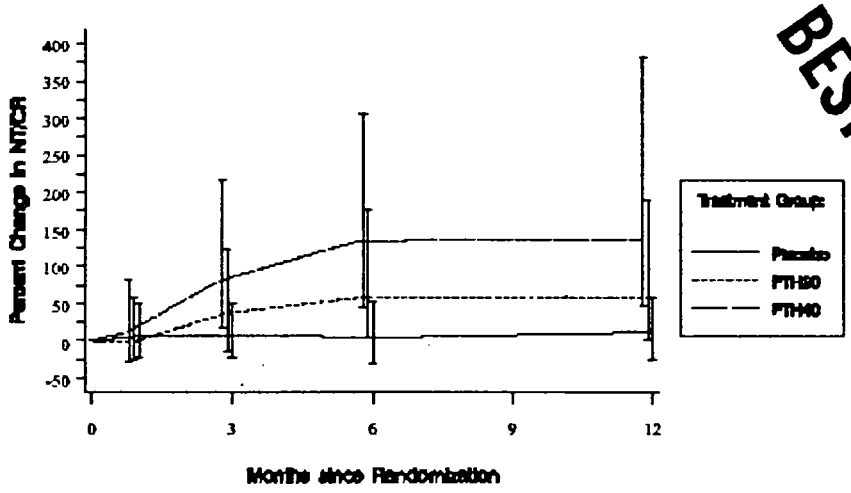
Median % changes from baseline in serum BSAP levels, by treatment group, are shown in the next figure:



Similar results were observed for the markers of bone resorption. Urinary NTX/Cr ratios increased significantly in both LY333334 treatment groups over baseline, after Month 1 ($p<0.001$), whereas there were small and inconsistent increases over baseline in the placebo groups, some of which were statistically significant ($p<0.05$). Statistically significant differences between each LY333334 treatment group and placebo were found at all time points except the first month. At study endpoint, the p-value was <0.02 for the 20 µg group vs placebo, and p-value was <0.001 for the 40 µg group vs placebo. At month 12, the increases over baseline were 11.5% for placebo, 59.1% for the 20 µg group, and 136.2% for the 40 µg group. At endpoint, the corresponding increases over baseline were 26%, 54.7%, and 105.5%.

The differences between the two LY333334 treatment groups was statistically significant at all visits ($p<0.001$), except Month 1.

The changes in NTX/Cr over time are depicted in the following figure:

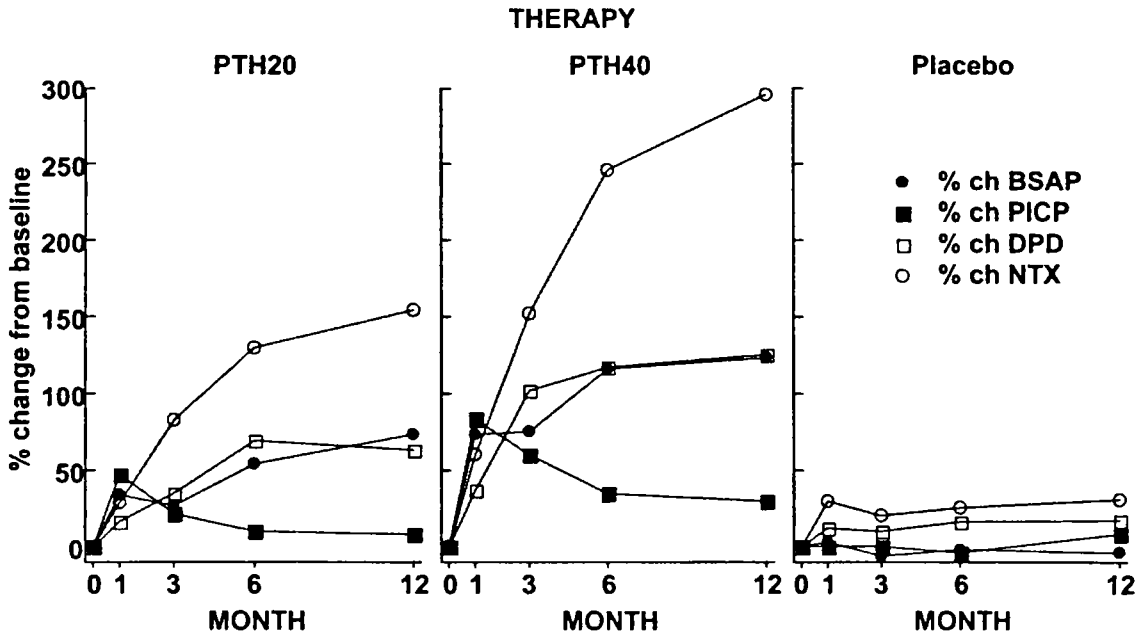


Analysis of urinary free deoxypyridinoline/Cr yielded results that were very similar to the NTX/Cr responses. Whereas levels of this resorption marker were increased only at endpoint in the placebo group (+14.5%), there were statistically significant increases over baseline in both LY333334 treatment groups at all time points after Month 1, including endpoint ($p<0.001$). For the 20 µg group, urinary deoxypyridinoline/Cr increased to 40.0% above baseline at Month 6 and was 30.4% over baseline at endpoint. The 40 µg group increased to 74.2% above baseline at Month 12, and to 50.0% over baseline at endpoint.

The difference between the 20 µg LY333334 group and placebo was significant at Months 3, 6, and 12 ($p<0.001$). The difference between the 40 µg group and placebo was significant at all time points after baseline, including endpoint ($p<0.001$). Consistent with results for the other markers, the difference between the two LY333334 groups was significant at all time points after baseline, including endpoint ($p<0.03$).

The responses of all four markers to each dose of LY333334 and to placebo are summarized in the next graph (courtesy of Joy Mele, Biometrics reviewer):

Statistical Reviewer's graph of biochemical markers — Study GHAC



For 1,25-dihydroxyvitamin D, the % change from baseline in the placebo group was not statistically significant at any scheduled time point. In contrast, for the LY333334 20 µg group, there were statistically significant increases in levels of the vitamin at all scheduled time points (i.e., months 1-12, $p<0.05$); and for the 40-µg group the increases were statistically significant at all scheduled visits except for Month 6. In both LY333334 treatment groups, the increase in levels of 1,25-dihydroxyvitamin D was rapid, reaching peak concentrations (27.0% above baseline in the 20 µg group and 24.0% in the 40-µg group) at Month 1. At endpoint, 1, 25- dihydroxyvitamin D concentrations in the 20 µg and 40 µg groups were 3.3% and 5.9% greater than baseline, respectively. These increases were not statistically significant.

The % change for the 20-µg group was statistically significantly greater than for placebo at Months 1 - 12 ($p<0.05$). With the exception of Month 6 ($p=0.059$), this was also true for the 40 µg group. There were no statistically significant differences between the two LY333334 treatment groups at any time point

Comments: The changes in biochemical markers of bone formation and resorption are consistent with the known anabolic actions of intermittently administered PTH 1-34, described in detail in the clinical pharmacology section of this review. In GHAC, there was evidence for enhanced pharmacodynamic responses as the dose of LY333334 increased from 20 to 40 µg/day. Also consistent with the known action of PTH and the overall

regulation of mineral metabolism, the increased bone formation was coupled with, and followed by, increased resorption. This was demonstrated by the increases in urinary NTX and deoxypyridinoline, both of which followed the elevations in formation markers. Finally, the elevations in all four markers declined following discontinuation of the drug, due to premature termination of the study. At study endpoint, which was several weeks after the last dose of LY333334, only PICP had returned to baseline levels.

The effects of LY333334 on levels of 1,25- dihydroxyvitamin D are consistent with the known action of parathyroid hormone on renal 1α-hydroxylase activity.

B.1.8.5.5 Other efficacy outcomes of GHAC

Other efficacy outcomes included population pk-pd evaluations, peripheral quantitative computed tomographic (pQCT) studies of the forearm, and health outcomes (including health-related quality of life indices). The last analysis does not appear in the proposed label and will therefore be reviewed only briefly here. Histomorphometric analysis of iliac crest bone biopsies are presented by the sponsor as part of the safety analysis. However, the results of these studies will be reviewed in this section.

The overall population pk-pd analysis has been reviewed by biopharmaceutics. In addition, an overview of this evaluation across all the clinical studies has been presented in this medical review, in **Section III** above.

The sponsor's population analysis of postmenopausal women in GHAC indicated that the pharmacokinetics of LY333334 are not meaningfully affected by age, body weight (over a broad range), alcohol consumption, smoking, or injection site (abdominal wall or thigh). There was no significant association between renal or hepatic function and clearance of LY333334, although patients with serious liver and kidney dysfunction were excluded from the trials.

The apparent volume of distribution (V/F) increased directly with body weight. V/F ranged from 66.6 to 199 L for typical patients weighing from 39.5 kg to 120 kg (the population minimum and maximum). When normalized for body weight, however, V/F was similar across the range of weights (approximately 1.7 L/kg). The effect of body weight on the volume of distribution of LY333334 does not significantly affect total systemic exposure (AUC), but it may affect C_{max}. The predicted peak serum concentration following injection of 20 μg LY333334 into the abdominal wall for a typical 39.5 kg patient is 236.6 pg/mL; for a typical patient weighing 120 kg, the predicted C_{max} is 90.1 pg/mL. These differences in peak concentration of the drug were not considered by the sponsor to be clinically significant, in terms of calcium responses and total systemic exposure. Five patients in the 20 μg group had a single measurement of LY333334 ranging

from 924 pg/mL to 3233 pg/mL. None of these patients had an elevated serum calcium concentration reported at any visit throughout the study, and none had a drug-related serious AE. Only 1 of these 5 patients had an elevated urine calcium excretion. In the entire study, there were 15 episodes (in 13 patients) in which the LY333334 level was > 600 pg/ml. As predicted, peak drug levels were higher in the 40 µg group.

Comments: As shown in the sponsor's Figure GHAC.11.27 (not reproduced here), many patients had post-dose PTH maxima that were between 300 and 600 pg/ml (at least twice as many in the 40 µg group, compared to the 20 µg group). However, these elevations of PTH were transient and had no obvious adverse effects on serum or urinary calcium levels. If there are any short- or long-term safety consequences of these transient elevations of PTH, they are unlikely to be mediated through effects on mineral metabolism.

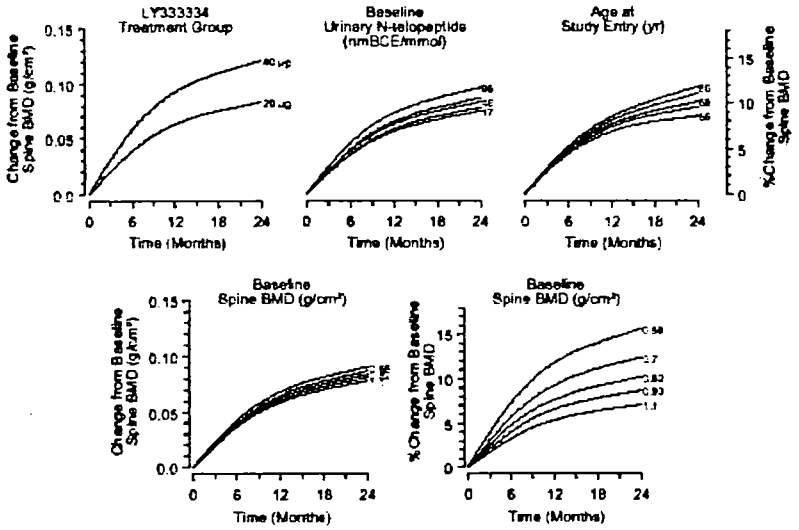
The volume of distribution of LY333334 was also found to correlate with injection site; V/F was about 21% higher after injection in the thigh, as compared to the abdomen. However, the BMD and biochemical bone marker responses did not differ according to injection site.

Population pharmacodynamic modeling for BMD and biochemical bone turnover markers revealed the following:

The lumbar spine BMD responses to LY333334 were greater in patients with lower lumbar spine BMD and/or higher urinary NTX at baseline. Older women also had greater lumbar spine BMD increases in response to LY333334, compared to younger women.

At the femoral neck, BMD increases in response to LY333334 were greatest in older patients with average body weight and high baseline urinary NTX excretion.

To illustrate the magnitude of these effects in the pharmacodynamic models, I have reproduced the sponsor's graphs of lumbar spine BMD responses in the next figure:



Pharmacodynamic modeling of biochemical markers of bone turnover revealed the following:

Baseline PICP and BSAP (formation markers) responded more rapidly to LY333334 than did the resorption markers NTX and DPD. However, rapid and substantial responses were found in all 4 markers. There was a near dose-proportional effect of LY333334 on all 4 biochemical markers.

Higher baseline levels of the formation markers were associated with a greater response of all 4 markers to LY333334.

The magnitude of the PICP increase at 1 month was a better predictor of the change in lumbar spine or femoral neck BMD at 19 months than were responses of other biochemical markers. The PICP increase was also a better predictor of the BMD response than was the dose of the drug.

Comments: The sponsor's population pharmacodynamic analyses were thorough and convincing. The results of these evaluations are not surprising, given the known pharmacological action of exogenous PTH. It is important to emphasize that no analysis identified specific baseline characteristics, or outer limits of values, that precluded efficacy of the drug (see, for example, lumbar spine BMD response curves in the figure above). This conclusion is essentially the same as in the analysis of subgroups.

Peripheral Quantitative Computed Tomography (pQCT)

The sponsor conducted an exploratory study that used pQCT technology to evaluate the treatment-related changes in the forearm (proximal and distal radius), skeletal sites that have both cortical and trabecular bone. This investigation was conducted at one study location. QCT has the ability to determine changes in density of cortical and cancellous bone compartments, volumetric bone mineral density (BMDv), and changes in trabecular microarchitecture.

In this sub-study, the sponsor used pQCT to measure cortical and cancellous bone geometry, mass, and volumetric bone mineral density at the proximal and distal radius. At the proximal radius, the data were used to predict biomechanical indices (axial and polar momentum of inertia, and torsional and flexural strength); these calculations could not be performed for the distal radius because the mathematical models could not be fitted to the geometry of that region.

Baseline measurements were performed at the distal, but not proximal, radius. For the latter, data are available for about100 patients at Months 12 and 24 and at study closeout. Results for the proximal were as differences between treatment groups at endpoint. For the distal radius, QCT measurements are available for about 50 patients at baseline, at Months 3, 6, 12, and 18, and at study closeout. Results for the distal radius were analyzed as change from baseline to endpoint.

At the proximal radius the sponsor measured the following primary pQCT indices: total bone area, total bone mass, total BMD, cortical BMDv , cortical area, cortical bone mass, cortical thickness, and periosteal and endosteal circumferences. Comparisons were made between treatment groups.

Results:

At the proximal radius, there were statistically significant treatment-related increases or trends in bone area, achieved by enlarging the periosteal, and (to a lesser degree) endosteal circumference without changing the bone mineral content. The periosteal circumference was about 5% higher in the 40 µg group, relative to placebo (p=0.005). There was a trend towards higher periosteal circumference in the 20 µg group, but the difference was not significant. There was a numerical increase in endosteal circumference (about 8.7% over placebo, p=0.051). There was an increase in total bone area, as a result of the increase in periosteal circumference in the 40 µg group, relative to placebo (p=0.006). Since the cortical bone mass (mineral content, BMC) did not change, the increased bone area resulted in a trend toward decreased cortical bone density in the 40 µg group (930.31 mg/cc, compared with placebo (961.78 mg/cc, p=0.054). In the 20 µg group, the increases over placebo were smaller in magnitude and were not statistically significant for any of these QCT parameters.

These changes resulted in increases in the computed biomechanical indices, axial and polar momentum of inertia and torsional strength, relative to placebo. The differences were statistically significant only in the 40 µg group, with positive trends in the 20 µg group.

At the distal radius, the number of patients was too small (there were 14-16 patients/treatment group) to show statistically significant treatment effects.

Comments: QCT technology, which is a topic of ongoing research and development, can provide a wealth of information about bone architecture in various disease states. In addition, this approach has the potential to provide insights into the action of drugs used to treat osteoporosis. This sub-study was too small, and probably inadequately designed, to provide definitive conclusions regarding the action of LY333334 on bone microarchitecture. The use of p-values in these multiple statistical comparisons is inappropriate. Limitations of this pilot study notwithstanding, the data are interesting and potentially important. For example, the increase in bone area without a corresponding enhancement in local BMC may explain the observed decrease in BMD at the distal 1/3 radius in GHAC (described above). It is possible that LY333334 may augment the mechanical strength of long bones in part by increasing the growth of cortical bone. This results in larger bone circumference and area, with consequent increases in strength. The positive effects of LY333334 on cortical bone are reassuring. It will be of great interest to study the effects of exogenous PTH using QCT at various skeletal sites. Such studies will provide additional insights into the long-term effects of the peptide on trabecular architecture. Given the unique action of PTH on bone anabolism, it is hoped that larger and longer studies of the effects of LY333334 on multiple indices of architecture, mass, and strength will be conducted.

Bone Histomorphometry:

Bone histology and histomorphometry were performed on paired trans-iliac biopsies (following double tetracycline labeling) from a subset of patients. The analysis concentrated on bone safety and evaluation of bone formation and resorption, mineralization, and trabecular architecture. The primary planned comparisons were the baseline to 1-year and 2-year changes in each treatment group. The sponsor intended that half of the patients who had a baseline bone biopsy would be randomly assigned to have a repeat biopsy at Month 12, and the other half, at Month 24. Approximately 120 patients were expected to enroll, in order to have at least 10 evaluable, paired biopsies per treatment group at each time point. All biopsies were analyzed in the laboratory of [redacted]. Complete details and standard histomorphometric parameters are presented in Appendix 16.1.12 of the submission.

The numbers of paired biopsies for each treatment group are presented in the following table:

	Placebo	PTH20	PTH40	Total
Number of baseline biopsies (number evaluable)	37 (31)	31 (26)	34 (30)	102 (87)
Number of 12-month biopsies (number evaluable)	8 (8)	7 (7)	6 (6)	21 (21)
Number of evaluable paired biopsies (baseline and 12 months)[a]	7	6	6	19
Number of biopsies at final study visit (number evaluable)	16 (14)	14 (13)	10 (9)	40 (36)
Number of evaluable paired biopsies (baseline and final study visit)[a]	12	11	8	31
Patients who did not have a follow-up biopsy performed	13	10	18	41

Abbreviations: PTH20 = LY333334 20 µg/day; PTH40 = LY333334 40 µg/day.

[a] Depending on variable analyzed, the number of biopsies suitable for analysis may be less than the nominal number of evaluable biopsies.

All bone biopsies were examined for the presence of osteomalacia, marrow fibrosis, cellular toxicity, woven bone, or any other abnormality. Histomorphometric indices were determined. These included structural, surface-based, and dynamic indices[7]. Results of these studies are presented in Tables GHAC.12.60-12.62 of the NDA submission.

No histological abnormalities were observed in the baseline biopsies. There was no woven bone, osteomalacia, or other histological abnormality observed in the postbaseline biopsies in any treatment group.

One biopsy in each of the PBO and 20 µg groups showed tunneling resorption at Month 12. There was significantly more tunneling resorption and bone marrow fibrosis in the 40 µg group, compared with PBO and the 20µg group. Tunneling resorption was observed in 5 (33%) patients in the 40 µg group: 4 (67%) at Month 12 and 1 (11%) at the final visit. A small amount of bone marrow fibrosis was observed in a total of 5 (33%) patients in the 40 µg group: 2 (33%) at Month 12 and 3 (33%) at the final visit.

The sponsor reports that there were no statistically significant histomorphometric differences ($p<0.05$) or trends ($p<0.10$) among the three treatment groups at Month 12. The exception to this was an increase in cortical porosity in the 40 µg treatment group, but not the 20 µg group. It should be noted that there were very few evaluable biopsies at this time point.

[7] Structural indices: trabecular bone and marrow star volumes, cortical thickness and porosity, osteoid thickness, and wall thickness. Surface-based indices: active eroded surface, total erosion surface, osteoid surface, and mineralizing surface. Dynamic indices: bone formation and resorption rates, mineral apposition rate, mineralization lag time and activation frequency.

At the final study visit, there was no longer an increase in cortical porosity in the 40 μg group. However, there were statistically significant increases ($p<0.05$) in trabecular bone volume, mean orthogonal intercept length, and mineral apposition rate, compared with placebo. Wall thickness and preosteoclast erosion depth showed trends to increase ($p<0.10$), while osteoclast erosion depth showed a trend to decrease, compared with placebo. The treatment effects were generally dose-related, with 40 μg treatment group showing greater effects than the 20 μg group. However, the sample size was not large enough to distinguish between the two LY333334 treatment groups.

In summary, no significant histological safety concerns (e.g., woven bone or osteomalacia) were identified by bone biopsy. There were no adverse histological effects observed in the 20 μg treatment group. There appeared to be a transient increase in cortical and trabecular remodeling in the 40 μg group that was seen at Month 12, but not at the final visit. The other, apparently dose-dependent histomorphometric effects that were seen at the final study visit are consistent with the anabolic effects of LY333334 on bone.

Comments: The sponsor applied summary statistics in the analysis of multiple histomorphometric parameters. Given the large number of comparisons, the use of p-values is inappropriate in the absence of pre-specified hypotheses and corrections for multiple comparisons. The number of evaluable paired biopsies was probably too small to yield very robust data regarding bone histomorphometry. Note that not all histomorphometric parameters could be measured in each of the evaluable paired biopsies. Nonetheless, there were several changes that were indicative of treatment-related anabolic effects on bone.

There were 37 evaluable biopsies from patients following at least 12 months of treatment. These showed no evidence of abnormal bone. Although this is not a large sample size, the data give some reassurance that LY333334 treatment is not associated with osteomalacia, woven bone, or other obvious pathology.

Health-related quality of life:

The sponsor studied the effects of treatment with LY333334, compared to placebo, on health-related quality of life (HRQOL), using a variety of instruments, with multiple endpoints. The results of this analysis are presented in detail in the NDA. As these results do not appear in labeling claims for LY333334, a detailed review will not be presented here. The following are a few brief comments on this analysis.

A combination of HRQOL questionnaires was administered to patients at several selected study sites in countries where adequate language translations an validations were available. The instruments were: the Nottingham Health Profile (NHP), the EQ-5D (formerly EuroQoL), the McMaster Health Utilities Index (MHUI), the Osteoporosis Assessment Questionnaire (OPAQ), the European Foundation for Osteoporosis Quality-of-Life Instrument (QUALEFFO, formerly EFFO), and the Attitude Index. These were administered to patients at baseline, Months 12 and 24, and at Early Discontinuation or the study closeout visit.

For the generic instruments, EQ-5D and NHP, patients' scores appeared to be stable over the treatment period, with statistically significant differences found only in the physical mobility sub-score of the NHP (out of 6 sub-scales). Here, the 40 µg group had a small deterioration, with slight improvement in the 20 µg group. The difference between the two LY333334 groups was statistically significant; however, with one exception, no other within- or between-group comparisons approached statistical significance, and overall the scores were quite stable over the 19 months. Of note, the exception was the overall pain score, which improved significantly over time in the placebo group, with no other significant comparative changes present. For EQ-5D, there were no relevant statistically significant within- or between-group changes for any scale, including Mobility and Pain/Discomfort.

For the MHUI, with one exception, there were also no statistically significant within- or between-group differences for any of the 6 attributes. The exception was the Sensation attribute, in which there was a significant difference between placebo and the 40 µg group. In this case, the LY333334 group score was lower, indicating a lower QOL. Of interest, in the two attributes Mobility and Pain, there were no significant differences in any of the group comparisons.

For the QUALEFFO (European Foundation for Osteoporosis Quality of Life) assessment, an instrument developed for patients with vertebral osteoporosis, the results were similar to the above three evaluations. There were no improvements (either statistically significant or as numerical trends) in the LY333334 groups, whether compared to placebo or to baseline. Of note, three of the five domains were Pain, Daily Activity, and Mobility. The other two were General Health and Mental Health.

Finally, in OPAQ (Osteoporosis Assessment Questionnaire), another osteoporosis-specific quality of life instrument, there were again no significant between-group differences for any domain, with the exception of the emotional status dimension. For this outcome, there was a worsening in mean change from baseline to endpoint in the 40 µg group compared to the 20 µg group. There were no other statistically significant differences (overall or pairwise) among the three treatment groups. There were small, statistically significant, within-group changes for some of the domains, with some worsening in three areas in the 40 µg group and some (equally small) beneficial changes in the 20 µg group.

The sponsor also analyzed changes in the HRQOL data in the subgroup of patients who experienced an incident vertebral fracture during the trial. This was designed to test the impact of fractures on HRQOL. The number of patients with fractures who filled out questionnaires was too small for statistical comparisons in two of the health-related instruments. Of the remaining three instruments, there were statistically significant differences (between patients with and without fractures) in the OPAQ emotional status scale and the OPAQ physical function scale. Of note, there were no differences between these two subgroups in pain or mobility ratings in the other two scales (NHP and MHUI).

Comments: The results of these quality of life investigations failed to disclose meaningful improvements as a result of LY333334 treatment, even in osteoporosis-related indices. The sponsor states that the trial period (about 20 months) was too short to assess changes in health-related quality of life, mainly because there was insufficient time to demonstrate a decline in quality of life in the placebo group. Therefore, the effects of LY333334 on these parameters may have been underestimated.

There is no basis on which to accept this opinion. Setting aside the failure of the statistical approach to account for multiplicity of endpoints, as well as the inappropriate use of the fracture subgroup (*vide supra*), there is not the least indication of improvement in osteoporosis-related quality of life outcomes (such as pain or mobility) in the results derived from the use of any of five separate analytical instruments. It is worth bearing this in mind when

GHAC convincingly demonstrated a treatment-related improvement in the proportion of patients with vertebral and non-vertebral fractures, as well as impressive increases in BMD. In my opinion, efficacy claims for this (or any other) drug are best confined to the realm of hard scientific data. Extending these clear and quantifiable outcomes to quality of life issues, including back pain, is potentially misleading, given the inadequacy of the methodology, the inappropriate use of p-values in *post hoc* statistical analyses, and the improper definition of analytical patient subgroups. These issues are discussed further in the summaries and in the review of the proposed labeling for Forteo (LY333334).

B.1.9 Summary of Efficacy for GHAC

GHAC was a pivotal, phase 3, randomized placebo-controlled multicenter trial of the safety and efficacy of LY333334 in the treatment of postmenopausal osteoporosis. In terms of size, endpoints, duration, and

ultimate marketing, GHAC was the most important trial in the clinical development program for LY333334.

The trial enrolled 1637 women, mean age 69 years (range 30-85 years), who had been postmenopausal for at least 5 years prior to randomization. Each patient had at least one moderate or 2 mild atraumatic vertebral fractures at baseline. For patients with fewer than 2 moderate fractures, the spine or hip BMD T-score had to be –1 or lower. Women were excluded if they had illnesses affecting bone or mineral metabolism (including Paget's Disease, hyper- or hypoparathyroidism, elevated endogenous PTH 1-84, and other diseases affecting bone), recent history of urolithiasis, impaired liver or renal function, or if they had taken drugs affecting bone or mineral metabolism within 2 years of enrollment. Patients were randomly assigned (1:1:1) to placebo, LY333334 20 μg/day, or LY333334 40 μg/day. All patients received adequate calcium + vitamin D supplementation (1000mg elemental calcium + 400-1200 IU of vitamin D). The study was originally planned to run for 3 years of double-blind treatment, but was prematurely interrupted (median treatment time 19 months) due to the finding of osteosarcomas in rats treated long-term with LY333334.

The baseline characteristics of the trial population were similar to those of patients in other studies of postmenopausal osteoporosis and essentially the same across all 3 arms of the study. The mean lumbar spine BMD T-score at baseline was –2.6 in the three treatment groups. The presence of prevalent baseline vertebral fractures undoubtedly ensured a reasonably high incident fracture rate during the trial. The overall retention rate was about 80% (until the trial was prematurely terminated by the sponsor), which is very much in keeping with results of most osteoporosis trials and certainly consistent with reliable analyses of endpoints. There were no significant differences in dropout rates between the 20 μg group and placebo; however, there was a small, but significant increase in discontinuation rate due to an adverse event in the 40 μg group, compared to placebo. The median exposure to the drug was 19 months. About 70% of patients in all 3 groups completed more than 17 months of treatment, and 82% completed more than 15 months.

Despite the premature termination of the study, the sponsor was able to meet the primary efficacy goal, as well as nearly all the secondary outcomes that are important for osteoporosis treatment trials. For the primary efficacy and lumbar spine BMD endpoints, the effects of 19 months of treatment with this new anabolic agent equaled or exceeded those that have resulted from 36-48 months of exposure to any known anti-resorptive drug. Although the pharmacodynamic responses (i.e., changes in levels of biochemical markers and increases in BMD) to 40 μg/day exceeded those of the 20 μg, the two doses were equal in anti-fracture efficacy. This finding was of particular importance, in view of the increased safety/tolerability

concerns that are associated with the higher dose of the drug. Thus GHAC, together with the extensive phase 2 studies, established 20 μg as the optimal daily dose of LY333334. However, the trial did not identify the ideal duration of treatment with LY333334, for reasons beyond the sponsor's control. This is unfortunate: judging from the trajectory of the BMD accumulation curves, one could reasonably anticipate substantial increases in bone mass with further treatment. It is quite likely that these further increases in BMD would translate into enhanced clinical benefits, given the anti-fracture efficacy of the drug during the initial 19 months and the necessary time interval between pharmacodynamic actions and clinical effects on bone strength. The potential benefits of adding an anti-resorptive agent during or after treatment with LY333334 remain unexplored. As noted above, we do not know the effects of giving the 20 μg dose less frequently (e.g., thrice weekly).

The study had one primary objective and 10 secondary objectives. The primary objective was *"to demonstrate a reduction in the proportion of patients with new vertebral fractures following 3-year treatment with 20 and 40 μgLY333334 plus calcium and vitamin D compared with calcium and vitamin D alone."* This outcome was clearly achieved. Of the 1636 women who entered the trial, 1326 (81%) had adequate baseline and follow-up spine radiographs. By study end, 105 patients had one or more new vertebral fractures, 64 (14.3%) in the placebo group, 22 (5.0%) in the LY333334 20 μg group, and 19 (4.4%) in the 40 μg group ($p<0.001$ for either LY333334 group vs placebo). This yields relative risk reductions of 65% for the 20 μg treatment group and 69% for the 40 μg group. The absolute risk reductions in these two LY333334 dose groups were 9.3% and 9.9%, respectively. In similar studies, the relative risks of suffering a new vertebral fracture were reduced by 47% with alendronate, 41% with risedronate, and 30% with raloxifene. At the time of this writing, nasal salmon calcitonin has not demonstrated consistent fracture reduction efficacy, and there are no data from randomized prospective trials of the fracture-prevention efficacy of estrogen. It is worth noting that the treatment periods for all the comparison drugs were substantially longer than 19 months.

The sponsor conducted further analyses of vertebral fractures in GHAC. Although these additional analyses were not pre-specified, they suggest that the benefits of LY333334 may extend beyond reduction in the proportion of patients with a new vertebral fracture. As part of the protocol for identification of morphometric fractures, the investigators counted the number of new fractures for each patient. In addition, the severity of the fractures was scored according to percent height reduction of each evaluable vertebral body. This analysis showed that there was a substantial reduction in the proportion of patients with multiple new vertebral fractures [22 (4.9%) in placebo, 5 (1.1%) in the 20 μg group, and 3 (0.7%) in the 40 μg

group]. The analytical protocol also included scoring of each evaluable vertebral body for % reduction in height. The reductions were classified as Grades 1, 2, and 3, which were termed "mild," "moderate," or "severe," according to the published protocol (Genant, 1993). Based on this classification, there were substantial reductions in the proportions of patients with "new moderate + severe" and "new severe" vertebral fractures in both LY333334 dose groups, compared to placebo. New moderate + severe fractures occurred in 42 (9.4%) placebo patients, in 4 (0.9%) patients in the 20 μg group, and in 9 (2.1%) patients in the 40 μg group. For new severe fractures, the results were 14 (3.1%) in placebo, none in the 20 μg group, and 3 (0.7%) in the 40 μg group.

Because the majority of morphometric vertebral fractures are clinically silent, it is difficult to evaluate the overall direct clinical impact of these data, taken alone. The presence of vertebral fractures is highly predictive of the occurrence of subsequent vertebral fractures, some of which will be clinically symptomatic. Unfortunately, the sponsor did not include an analysis of clinical vertebral fractures in this application. These are fractures that usually present as back pain and are confirmed radiologically. The incidence of these has been low in past clinical trials, and it has been difficult to power trials to examine this endpoint specifically. For example, in the Vertebral Fracture Study of FIT, which enrolled patients with a prevalent vertebral fracture (similar to GHAC), the clinical fracture rates were 5.0% in the placebo group and 2.3% in the alendronate group over 3 years. These rates are substantially higher than in patients without a baseline vertebral fracture [e.g., in the clinical fracture arm of FIT (patients without a vertebral fracture) the rates were 1 [illegible]% in placebo vs 1.04% in the alendronate group]. Thus in the former, [illegible]-risk group, there were similar relative risk reductions for clinical and morphometric vertebral fractures (in the 45-55% range). Whether increases in the size and duration of GHAC would have permitted a comparable analysis is not clear. However, as noted above, in most (42 out of 64) of the placebo patients with new fractures, the fractures were moderate or severe in grade. In contrast, of the 22 patients who had a new fracture while taking LY333334 20 μg/day, only 4 had a moderate fracture, and none had a severe fracture. This means that 18 of the 22 patients who experienced vertebral fractures while taking 20 μg LY333334 had Grade 1 fractures.

These results strongly suggest that LY333334 is capable of preventing multiple new vertebral fractures, as well as reducing the severity of these fractures. According to the sponsor, other outcomes of the study indicate that these effects translate into diminution of height loss [illegible] This interpretation forms the basis of two proposed labeling claims for LY333334.

As noted above, this subgroup analysis flawed. The definition of a subgroup based on events that occur during a trial is improper. In addition, the analysis was not pre-specified. A description of how this fallacy led to irreproducible conclusions in two alendronate trials (in which the first data set yielded even more robust treatment differences) has also been presented in this review. Other drugs (alendronate, risedronate) have demonstrated beneficial effects on height loss in clinical trials, although the treatment-related differences have been small. It should be remembered that morphometric fractures are a dichotomous variable, while height changes are continuous. For example, a patient with a 20% loss of vertebral height would be classified as having suffered a fracture, but one with a 19% loss would not. Given all the benefits of LY333334 on the lumbar spine, it is remains unclear why a favorable effect on height should not have been observed in the trial population as a whole.

Finally, the sponsor's claims that treatment with LY333334 reduces the proportion of patients with multiple vertebral fractures as well as the severity of these fractures cannot be accepted on the basis of this trial. There is no question that these outcomes occurred during the trial. The problem is that none of these results was pre-specified as a primary or secondary outcome. In the course of counting fractures, according to the protocol, the grade (alternatively called "severity") and number of fractures were recorded. It is my understanding that it is not necessary to record either of these parameters in order to establish whether a particular individual suffered at least one new vertebral fracture. Accordingly, these are best described as protocol- or method-derived results that were not necessary to answer the question posed by the primary hypothesis. In claiming these as efficacy outcomes, the sponsor has elevated method-derived data to the level of efficacy outcome variables. I have already discussed the associated issues of multiplicity and inappropriate use of p-values in this context.

Thus, the primary efficacy outcome was clearly achieved. Although many related events occurred in this trial, there is no basis for describing them as pre-specified efficacy outcomes or assigning p-values to their "statistical significance."

Secondary endpoints

A key secondary endpoint was the proportion of patients with new non-vertebral fractures and with new vertebral and non-vertebral fractures combined. Each LY333334 treatment group demonstrated a statistically significantly lower proportion of patients with new non-vertebral fractures, whether examined as total fractures or as non-traumatic fractures. However, the treatment-related effects were far less robust than for morphometric vertebral fractures. For total non-vertebral fractures, there were 6.3% in the 20 μg and 5.8% in the 40 μg group, compared with 9.7% in the placebo group; relative risk reduction for the 20 μg and 40 μg LY333334 treatment groups, compared to placebo, was 35% and 40%, respectively ($p<0.05$). The absolute risk reduction for a non-vertebral fracture was 3.6% in the 20 μg group and 3.9% in the 40 μg group. There was no correction for multiplicity of outcomes in any of these comparisons.

Non-vertebral non-traumatic fractures were reduced by 53% in the 20 μg LY333334 group, and by 54% in the 40 μg group ($p<0.02$ for comparison of 20 μg group with PBO); however, this sub-analysis was not pre-specified as an efficacy outcome. The two active LY333334 treatment groups did not differ significantly in non-vertebral fracture risk reduction.

These results can be compared to those derived from studies of similar populations, in which the risk of non-vertebral fractures was reduced by 20

% with alendronate, 39% with risedronate, and 10% with raloxifene. The duration of these studies was three years.

This study lacked sufficient statistical power to detect treatment-related differences at individual non-vertebral sites, and the numbers of fractures at these sites were low. At each non-vertebral site, there were numerically fewer fractures in the treatment groups than in placebo. At the hip there were 4 fractures in placebo, 2 in the 20 μg group, and 3 in the 40 μg group. At the wrist, the corresponding numbers of fractures were 13, 7, and 10. Thus, despite the overall substantial and statistically significant efficacy in preventing non-vertebral fractures as a group, LY333334, 20 μg/day, prevented 2 hip and 6 radius fractures in 541 patients treated for about 19 months. Clearly, a longer, and probably larger, trial would have been needed to demonstrate efficacy at individual extra-vertebral sites.

For vertebral and non-vertebral fractures combined (another pre-specified secondary endpoint), the three treatment groups showed statistically significant differences in the proportions of patients with at least one fracture, compared to placebo. The 20 μg and 40 μg groups had relative fracture reduction rates of 51% ($p<0.001$) and 54% ($p<0.001$), respectively. Problems with interpretation of this hybrid endpoint are discussed above.

Although not specified as a secondary endpoint, the cumulative incidence of one or more new non-vertebral fractures was similar in the three treatment groups until about 12 months, when the protective effects of LY333334 became apparent (see Kaplan-Meier curves above). This is consistent with the necessary delay between pharmacodynamic action and fracture prevention.

LY333334 treatment increased BMD at all skeletal sites, except the ultradistal and distal radius. The increases were consistently greater at treatment end than at month 12, and were also greater in the 40 μg treatment group, compared to the 20 μg group. There were small increases in total body BMC in both LY333334 groups, whereas the placebo group lost total body BMC by study end. Consistent with the anabolic mechanism of the peptide, the speed and magnitude of increases in lumbar spine BMD exceeded those of any known anti-resorptive agent. However, the increases in BMD at other skeletal sites were generally no greater than have been seen following 12-24 months of treatment with bisphosphonates. Because GHAC was terminated prematurely, the increases in BMD following longer treatment periods are not known.

A responder analysis showed that nearly all patients treated with either dose of LY333334 gained spinal BMD. Most patients gained 5% over baseline. Further analyses identified no patient subgroups in which the

drug did not produce substantial and statistically significant increases in spinal BMD, relative to baseline and to placebo.

Changes in biochemical markers of bone formation and resorption were consistent with the anabolic action of LY333334. The pharmacodynamic responses to LY333334 40 μg were greater than observed in the 20 μg group; this dose-dependence was consistently observed during the clinical pharmacology studies. Also consistent with the known physiology of PTH, the increased bone formation (BSAP and PICP) was coupled to increased resorption, as shown by increases in urinary NTX and deoxypyridinoline. The increases in the two resorption markers followed the elevations in formation markers. Finally, the elevations in all four markers declined following discontinuation of the drug, due to premature termination of the study. At study endpoint, which was several weeks after the last dose of LY333334, only PICP had returned to baseline levels. The effects of LY333334 on levels of 1,25- dihydroxyvitamin D are consistent with the known action of parathyroid hormone on renal 1α-hydroxylase activity.

Extensive population-based pharmacokinetic-pharmacodynamic studies also failed to identify baseline characteristics that would necessitate LY333334 dose adjustments. This statement is qualified by the limits of renal and hepatic impairment that were present in the trial population. No pk-pd analysis identified specific baseline characteristics, or outer limits of baseline values (e.g., BMD), that precluded efficacy of LY333334.

Bone histology, carried out on a subset of patients, showed no evidence of abnormal bone following treatment with LY333334. Histomorphometric analysis of paired biopsy specimens showed evidence of anabolic action of the drug on bone.

As secondary endpoints, the sponsor also employed five independent QOL indicators. The results of this analysis failed to disclose any meaningful improvements as a result of LY333334 treatment, even in the two osteoporosis-related indices. No labeling claims are made for QOL improvements.

In conclusion:

- GHAC convincingly demonstrated a substantial treatment-related improvement in the proportion of patients with morphometric vertebral and pooled non-vertebral fractures, as well as impressive increases in spinal BMD and considerable increases in BMD at nearly all other skeletal sites. After 19 months of treatment, the reduction in risk of vertebral fractures, and the increases in spinal BMD, were greater than reported following longer treatment with any currently approved agent.

- There was no effect of LY333334 treatment on height loss in the study group as a whole.
- There was no effect of LY333334 treatment on five health-related quality of life indicators.
- Although the pharmacodynamic effects of 40 μg/day LY333334 exceeded those of the 20 μg/day dose, the fracture efficacies (both vertebral and non-vertebral) of the two doses were indistinguishable. Given the added safety/tolerability concerns of the higher dose, GHAC successfully established 20 μg/day as the indicated dose for all adult patients.

- Extensive population-based pharmacokinetic-pharmacodynamic modeling disclosed no population group or baseline characteristic that would preclude substantial and statistically significant efficacy of LY333334 in increasing lumbar spinal BMD. In addition, these studies have indicated that dose adjustments are not required on the basis of any baseline demographic or other characteristic, within the limits of the study.

- Histological analyses of iliac crest biopsies revealed no suggestion of abnormal bone in association with LY333334 treatment. There was evidence in support of an anabolic action of the drug on bone.

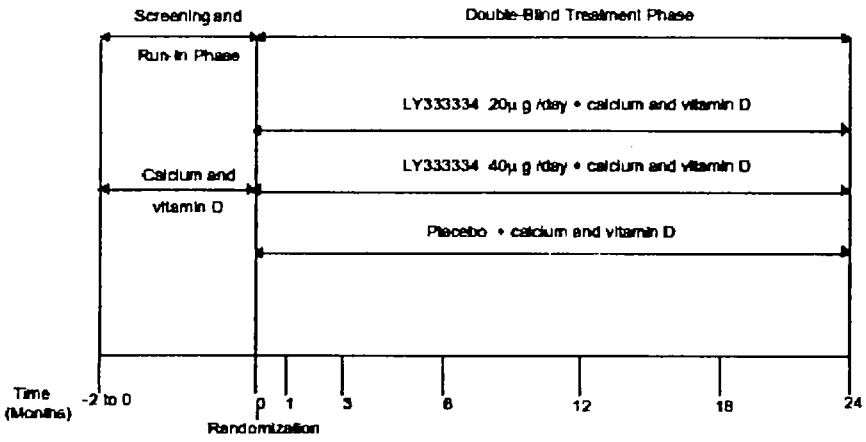
- A complete review of the safety of LY333334 in GHAC is included in the Integrated Safety Review, below. In general, no unanticipated serious treatment-emergent adverse events occurred during the study. However, as discussed in the safety review, there is a need for additional evaluation of cardiovascular responses to LY333334. In addition, in view of the dose-dependent occurrence of osteosarcomas in rodents treated with LY333334, there remains a need to establish a long-term monitoring mechanism post-approval. In this regard, the drug should be labeled to indicate that Paget's disease (which carries an increased risk for osteosarcoma) should be ruled out by history and appropriate screening prior to use.

B.2 Reviewer's Trial #2, sponsor's trial GHAJ

Effects of LY333334 in the treatment of men with osteoporosis

B.2.1 Design

GHAJ was a Phase 3, multicenter (37 study centers), double-blind, placebo-controlled, parallel, randomized study of the effects of LY333334 in 437 men with primary osteoporosis. The study was originally planned to include a 1-2 month calcium + vitamin D run-in phase, followed by a-year placebo-controlled treatment period. Patients were randomized to LY333334 20 µg, LY333334 40 µg, or placebo (1:1:1). All patients received calcium plus vitamin D supplementation. The trial design is depicted in the next figure:



B.2.2 Objectives

The primary objective was *"to demonstrate an increase in vertebral BMD in men with primary osteoporosis following 2-year treatment with LY333334 40 µg/day plus calcium and vitamin D or LY333334 20 µg/day plus calcium and vitamin D, compared with patients treated with calcium and vitamin D alone."*

The study had 8 secondary objectives. As stated by the sponsor, these were:

"To establish the effect of long-term treatment with LY333334 plus calcium and vitamin D, compared with calcium and vitamin D alone, on the following:
1) biochemical markers of bone format ion and resorption (serum procollagen I C-terminal propeptide [PICP], bone-specific alkaline phosphatase, urinary N-telopeptide, and urinary free deoxypyridinoline)
2) hip BMD
3) total body and radial (forearm) BMD
4) height (via ——— stadiometer or other suitable stadiometer)
5) clinical and biochemical safety parameters.
6) To quantify medical resources used by patients during the study so that a cost-effectiveness analysis could be performed.
7) To assess the impact of LY333334 on health-related quality of life in men with primary osteoporosis at study sites where translated and validated instruments are available.
8) To assess population pharmacokinetics of LY333334 in men."

B.2.3 Protocol

B.2.3.1 Populations

The trial population consisted of ambulatory men, aged 30-85 years, with primary osteoporosis. Primary osteoporosis is defined as bone loss that is either idiopathic or due to primary hypogonadism[8]. Idiopathic osteoporosis means that the bone loss is not due to hypogonadism or to other secondary causes, such as corticosteroid use or growth hormone deficiency.

Osteoporosis is defined in this trial as PA lumbar spine (L1-L4 or L2-L4) BMD or hip BMD T-score ≤ -2.0.

---

[8] The NDA protocol defined primary hypogonadism as low testosterone OR elevated gonadotropins. This definition might include patients with low testosterone and normal or low gonadotrophins, which would include cases of secondary hypogonadism. Other questions regarding establishment of the diagnosis of hypopituitarism were not addressed in the submission. These issues were clarified in correspondence with Lilly (May 25, 2001): "The patient's type of osteoporosis was further classified as either hypogonadal in patients with a low Free Testosterone (FT) or with elevated gonadotropins or in absence of other causes of their osteoporosis as idiopathic. Of the men who were classified as hypogonadal, 94.2% (196 out of 208) had a biochemical profile consistent with primary hypogonadism. The remaining 12 men who had both low FT as well as low gonadotropins could possibly be classified as having secondary hypogonadism. Patients who had a known Growth Hormone (GH) deficiency from any cause including previous pituitary surgery, pituitary tumor or pituitary radiation were not eligible for the study. Although section 3.4.3. Disease Diagnostic Criteria only deals with isolated pituitary failure (GH deficiency) the pituitary surgery, pituitary tumor and pituitary radiation most likely would classify these men as "pituitary failures" and would have led to exclusion from the study." (S.Zalani, Reg. Affairs Officer, Eli Lilly, May 25, 2001)

Other inclusion criteria:

- Absence of "severe or chronically disabling conditions other than osteoporosis (for example, uncontrolled diabetes or diabetes with significant renal, vascular, neurologic, or ophthalmic complications)."
- Absence of language barrier; "cooperative, expected to return for all follow-up procedures, and who gave informed consent before entering the study and after being informed of the medications and procedures used in this study."
- PA lumbar spine (L-1 to L-4 or L-2 to L-4) BMD or hip BMD at least 2.0 standard deviations below the average for young, healthy men (T-score).
- Intact L-2 to L-4 vertebrae, without artifacts, crush fractures, or other abnormalities which might interfere with the analysis of the PA lumbar spine BMD measurement.
- Normal or "clinically insignificant abnormal" laboratory values (baseline serum Calcium, PTH [1-84], and 24-hour urine calcium within normal limits; 25-hydroxyvitamin D between the lower limit of normal and three times the upper limit of normal.

Exclusion criteria:

- Any concurrent (or within 1 year prior to randomization) metabolic bone disorder other than osteoporosis (e.g., Paget's disease, renal osteodystrophy, or osteomalacia); secondary osteoporosis.
- Concurrent or recent (within 1 year prior to randomization) disease that can affect bone metabolism (e.g., hyperthyroidism, hyperparathyroidism, hypoparathyroidism).
- Currently suspected carcinoma, or history of carcinoma in the 5 years prior to randomization, except for excised skin lesions (e.g., basal or squamous cell carcinoma).
- Nephrolithiasis or urolithiasis within 2 years of randomization. Individuals with any history of nephro- or urolithiasis must have documented absence of active disease by an appropriate radiological evaluation (e.g. i.v.p. or supine radiograph) within 6 months of randomization.
- Concurrent or recent (within 1 year prior to randomization) sprue, inflammatory bowel disease, malabsorption syndrome. Presence of any indication of intestinal malabsorption of dietary calcium, e.g., low urinary calcium combined with elevated endogenous PTH level.
- Significantly impaired hepatic function. This was defined as: a single transaminase (ALT, AST, or GGT) > 3X ULN; or total bilirubin > 2.0 mg/dl; or multiple transaminase elevations of any magnitude which, in the opinion of the investigator, indicated significant hepatic impairment.
- Significantly impaired renal function. This was defined as: serum creatinine >2.0 mg/dL, or measured or calculated creatinine clearance which were indicative of significant renal impairment, in the opinion of the investigator.
- Current or recent (within 1 year) abuse of alcohol (defined in NDA) or drugs.
- Poor medical or psychiatric risk.

- Initiation or significant change in androgen or other anabolic steroid treatment within 6 months prior to randomization. Patients treated with stable doses of androgens or other anabolic steroids for at least 6 months prior to randomization could remain on their medication during the study.
- Treatment with calcitonin within 2 months of randomization. Treatment with progestins, estrogens, and estrogen antagonists within 6 months prior to randomization.
- Treatment with corticosteroids. This included systemic glucocorticoid equivalent of > 5 mg prednisone/day within 1 month of randomization, or for more than 30 days during the 12 months prior to randomization, or for more than 30 days in any 12-month period post-randomization. Topical corticosteroid use was limited, as described in the protocol. In addition, strict limitations were placed on use of inhaled corticosteroids. Doses > 840 µg/day beclomethasone, or equivalent, for more than 30 days in the year prior to randomization, or in any year post-randomization, were prohibited. Patients could not receive more than 4 intra-articular injections of equivalent of >40 mg triamcinolone within the year prior to randomization or in any 1-year period post-randomization. There was no limitation placed on use of ophthalmic and otic corticosteroids.
- Treatment with fluorides in 6 months prior to randomization or for more than 60 days in the 2 years prior to randomization. Fluoridated water and topical dental fluoride were permitted.
- Treatment with bisphosphonates in the 3 months prior to randomization, or for more than 60 days during the year prior to randomization. Intravenous bisphosphonates were not permitted within 2 years prior to randomization.
- Treatment with Vitamin D in doses > 50 000 units/week, or with any dose of vitamin D analog, during the 2 months prior to randomization.
- Use of coumarins, heparin, anticonvulsants (except benzodiazepines) were all limited (as specified in protocol) or completely precluded during the 6 months prior to randomization.
- Calcium- or aluminum-containing antacid use was limited, as specified in the protocol.

B.2.3.2 Discontinuations from trial

Patients could be discontinued from treatment for any of the reasons listed in protocol for GHAC (see review of GHAC above). The algorithm for reducing the dose of calcium supplementation and/or dose of the study drug on the basis of appearance of hypercalcemia and/or hypercalciuria (>350 mg/24 hours; or elevated urinary calcium/creatinine ratio) was also the same as in GHAC. The protocol also provided for LY333334 dose reductions in the event of repeated nausea, dizziness, or other adverse events.

As in GHAC, patients could be discontinued in the event of accelerated bone loss. The protocol specified discontinuation for lumbar spine BMD reductions from baseline of >7% in first year or >9% during the second year. Patients could be withdrawn if the BMD at the hip were reduced by 9% in the first year or by 11% in the second year.

Other possible reasons for discontinuation are presented in the NDA; these are the same as for GHAC. In the event of discontinuation, the protocol specified a battery of data that were to be collected. These included a physical examination with recording of height, weight, sitting blood pressure, and heart rate; assessment of habits (smoking, alcohol, caffeine); assessment of dietary calcium intake; collection of information regarding adverse events and concomitant medications; pre-specified laboratory assessments (hematology, clinical chemistry, urinalysis, serum calcium and albumin (4 to 6 hours post-dose); 25-hydroxyvitamin D; 24-hour urine calcium, creatinine, and phosphorus; and LY333334 antibodies).

Upon discontinuation, patients also had lateral thoracic and lumbar spine x-ray films, posterior-anterior (PA) lumbar spine BMD, hip BMD, and health-related QOL and resource utilization assessments. A subset of patients received additional laboratory assessments on discontinuation. These included measurements of levels of serum 1,25-dihydroxyvitamin D, biochemical markers of bone metabolism, fasting serum triglycerides, and serum LY333334. In addition, these patients received total body and radial BMD determinations.

B.2.4 Treatments, concomitant therapies, and schedule of events

B.2.4.1 Treatments

The treatments were essentially the same as in GHAC. During the run-in period, all patients were instructed to take daily calcium supplementation (approximately 1000 mg/day of elemental calcium), plus vitamin D (about 400 - 1200 IU/day) beginning at least 1 month prior to baseline and until the end of the treatment phase. As with GHAC, each patient was instructed in use of the injection device [redacted]. After screening and successful completion of the run-in phase, patients were randomly assigned to one of the three treatment arms. After review of proper use of the injection device, patients selected either the lower abdomen or the outer thigh for the injection site and used that area for injection for the remainder of the study. Injection sites were alternated from side to side each day. Patients could select the time of day for injection, without regard to meals. Injections were given at approximately the same time each day. The calcium and vitamin D supplements could be taken at any time of day.

B.2.4.2 Prior and concomitant medications

Excluded medications are indicated above. According to the sponsor, none of the prohibited medications were taken during the study, with a few exceptions.

B.2.4.3 Compliance

Individual investigators monitored patients' compliance with treatment. Patients returned injectable materials at Months 3, 6, 12, and Early Discontinuation or study closeout visit. Any remaining material was quantified and recorded. The number of used, partially used, and unused injection devices was also recorded. Any patient who missed > 50% of doses over two consecutive visit intervals and had participated in the study for *more* than 1 year post-randomization may have been discontinued. All patients with post-baseline data were part of the intent-to-treat analysis.

B.2.4.4 Schedule of events

The original protocol called for a 2-month run-in period, followed by two years of placebo-controlled treatment. As originally planned, the schedule of events and assessments are given in the following table:

Month Visit	-2 to 0 1	0 2	1 3	3 4	6 5	12 6	18 7	24 8	E D
Activity									
Informed consent document signed	X								
Patient number assigned	X								
Review of inclusion/exclusion criteria	X	X							
Medical history	X								
Physical examination	X					X		X	X
Sitting blood pressure, pulse	X	X	X	X	X	X	X	X	X
Concomitant medication reporting	X	X	X	X	X	X	X	X	X
Patient education regarding injections	X	X							
Record of adverse event reporting	X	X	X	X	X	X	X	X	X
Height[a]	X					X		X	X
Weight	X		X	X	X	X	X	X	X
Assessments of smoking, alcohol, caffeine	X					X		X	X
Assessment of dietary calcium	X							X	X
Patient assigned to treatment group		X							
Study material dispensed		X		X	X	X	X		
Laboratory Assessments									
TSH	X								
Estradiol, FSH, LH	X								
Morning total and free testosterone	X	X	X						
LY333334 antibodies	X					X		X	X
Hematology, clinical chemistry, urinalysis	X				X	X	X	X	X
Serum calcium, albumin (4-6 hours postdose)			X	X	X	X	X	X	X
25-hydroxyvitamin D	X					X		X	X
PTH(1-84)	X					X		X	X
24-hour urine calcium, creatinine, phosphorous	X		X		X	X		X	X
Biochemical markers of bone formation and resorption[b]		X	X	X	X	X	X	X	X
1,25-dihydroxyvitamin D		X	X	X	X	X	X	X	X
LY333334 serum concentrations			X	X	X	X	X	X	X
Technical Assessments									
Lateral thoracic and lumbar spinal x-ray films[c]	X								
Electrocardiogram	X								
PA lumbar spine BMD	X[d,e]			X	X	X	X	X	X
Hip BMD	X[d,e]					X		X	X
Total body and radial (forearm) BMD[f]	X[e]					X		X	X
Miscellaneous									
Medical resource utilization		X			X	X	X	X	X
Health-related quality of life		X			X	X	X	X	X

X = Performed during this time interval.

B.2.5 Efficacy outcomes

The primary efficacy outcome was % change from baseline in lumbar spine BMD following 2 years of treatment with LY333334, compared to placebo. The 8 secondary objectives have been listed above. The methodology for assessment of BMD at individual skeletal sites has been described in the review of GHAC, above. Lumbar spine BMD was to have been measured repeatedly, at several specific time points throughout the 24 months. As discussed below, because of the sponsor's decision to terminate the trial early, spine BMD was measured at

baseline, Months 3, 6, and 12, and Early Discontinuation or study closeout visit. Similarly, patients had hip BMD measurements at baseline, Month 12, and Early Discontinuation or study closeout visit. All patients contributed data to this analysis, using LOCF. It should be noted that not all patients had an assessment following 12 months of treatment. "Month 12 data" were based on both LOCF plus data derived from a scheduled visit following 12 months of therapy.

Total body and radius BMD determinations were assessed at baseline and at Month 12 and Early Discontinuation visit or study closeout.

The lumbar spine and hip BMD determinations were performed at screening and then twice during the run-in phase. The radius BMD determinations were also performed twice during the screening/run-in phase, following determination that the patient was eligible for entry into the study (on the basis of hip and spine BMD).

As for GHAC, the original design included two sub-studies (described below in the Statistics section). The sponsor presents data for spine BMD by sub-study and as overall results. Other BMD data are presented as overall results. All BMD data are shown as mean change and mean percent change from baseline.

The circumstances governing the early termination of this study, and the consequent changes in the data analysis, are discussed below. Similar to GHAC, the sponsor added a by-visit analysis in order to evaluate the time course of changes in BMD.

Comments: There is some confusion, throughout the GHAJ report, regarding the data set that is to be used for the BMD efficacy analyses: Visit 6 with LOCF or Visit 7 with LOCF. Apparently, the sponsor decided on Visit 6 with LOCF. As described below, this decision was arbitrary. Further comments on this issue appear in the sections below.

The sponsor presents BMD T-score cutoff values for men at the lumbar spine, femoral neck, and total hip, using both Lunar and Hologic densitometers:

	Hologic BMD (g/cm^2)	Lunar BMD (g/cm^2)
Lumbar Spine		
-2.0 SD	0.871	0.980
Femoral Neck		
-2.0 SD	0.759	0.830
Total Hip		
-2.0 SD	0.812	–

Abbreviations: BMD = bone mineral density; SD = standard deviations below the mean for young race- and gender-matched controls (North American Standard).

Biochemical markers of bone formation and resorption (serum PICP, BSAP, urinary N-telopeptide, and urinary free deoxypyridinoline) and 1,25-dihydroxyvitamin D were measured at baseline and at Months 1, 3, 6, and 12, and Early Discontinuation visit or study closeout. All samples were obtained in the morning and stored centrally [redacted]

Height was assessed at baseline (screening/run-in period) and at Month 12 and Early Discontinuation visit or study closeout. Measurements were performed with a stadiometer, as in GHAC.

Serum concentrations of immunoreactive LY333334 were determined at Months 1, 3, 6, and 12, and Early Discontinuation visit or study closeout. The methodology for these determinations has been described above in the clinical pharmacology section, and in the review of GHAC. Blood samples were collected at the individual study site, at specified time intervals, for periods up to 4 hours following injection of the drug. Blood samples were taken from randomly allocated patients according to a sampling schedule. As for GHAC, the overall sampling scheme was designed to gather population pk data. Thus a sparse sampling strategy was used to provide a few samples from many patients, rather than many samples from a few patients. The population pharmacokinetics of LY333334 were based on these data. LY333334 concentrations were assayed using an [redacted] directed at two sites in the human PTH 1-34 sequence. The sensitivity of this assay is 50 pg/ml. Comments on the limitations of this assay appear above, in the review of the clinical pharmacology section.

The sponsor used five health-related quality of life instruments, in different combinations. The maximum number of instruments at any one study site was three. Some study sites did not measure any QOL parameters. A listing of QOL instruments by country is provided in the NDA. The sponsor employed the same instruments as in GHAC, described above. These included three general health profiles [the Nottingham Health Profile (NHP), Part I of EQ-5D (formerly known as EuroQoL), and the McMaster Health Utilities Index (MHUI)] and two osteoporosis-specific instruments [the Osteoporosis Assessment Questionnaire (OPAQ) and the European Foundation for Osteoporosis Quality of Life Instrument (QUALEFFO).

The health-related QOL questionnaires were administered at baseline and at Months 6, 12, and Early Discontinuation visit or at study closeout.

Safety examinations were performed according to a pre-specified protocol, presented in detail in the Integrated Review of Safety.

Comments: The protocol does not include electrocardiograms, except at baseline. There are no electrocardiograms taken immediately after administration of LY333334. The protocol also does not include determinations of vital signs following administration of LY333334 at the study sites. In my opinion, these are significant omissions, especially in view of the potential for vasoactive responses to the peptide. Since the drug will presumably be marketed to elderly men and women, this issue should be addressed in a prospective study. In addition, potential interactions of LY333334 with digitalis, using EKG outcomes, should also be evaluated.

B.2.6 Statistical considerations

Patients in GHAJ were randomly assigned into one of the three treatment groups: LY333334 40-µg/day, LY333334 20-µg/day, or placebo. All patients received daily supplementation with calcium and vitamin D, as described. GHAJ was originally expected to have approximately 30 investigators, each enrolling about 6 to 12 patients. Investigators enrolling less than 6 patients were combined as a single investigator group, for analytical purposes. For the efficacy variables, GHAJ was analyzed both as a whole and within each of the 2 sub-studies. Allocation of sites into the sub-studies was done according to a prospective scheme, presented in the NDA. The scheme depended on allocation according to enrollment size, alternating in descending order.

All treatment comparisons used a 2-tailed test with a nominal significance level of 0.05. All efficacy analyses were done using an intent-to-treat (ITT) approach. This approach analyzed data by the groups to which patients were randomly allocated, whether or not the patients received assigned treatments, or did not otherwise follow the protocol. Only patients who had at least one post-baseline observation for the variable of interest were included in this analysis.

Reasons for discontinuation were compared between groups using Chi-square tests. The sponsor summarized baseline patient characteristics by treatment groups. To test for differences between treatment groups, the sponsor used an ANOVA model for continuous variables and Chi-square tests for categorical parameters. For both absolute and % change from baseline, between-treatment comparisons were performed at each visit and at endpoint. An ANOVA was used for comparisons that included the effects of treatment and investigator. The sponsor used pairwise comparisons for the three treatments groups using the least squares means without adjusting for multiple comparisons. According to the sponsor, this was due to the small number of treatments. Safety parameters of vital signs and laboratory values were evaluated similarly at each visit and at endpoint. The frequency and % of adverse events were presented for each treatment group and compared using a Cochran-Mantel-Haenszel test adjusting for investigator.

Comments regarding the use of multiple endpoints, endpoints that were not pre-specified, and the inappropriate use of certain analytical subgroups appear throughout this review, as well as in the Statistics Review.

The sponsor developed a population pharmacokinetic model, as described in the NDA, to characterize patient factors that may influence the disposition of LY333334 in osteoporotic men. Details are provided in the NDA. The model has been described in sections above as well as in the Biopharmaceutics Review. Patient factors that were assessed in the population pk analysis were essentially the same as for women in GHAC, with the addition of free testosterone levels.

A statistical power calculation is based on enrollment of 279 patients and anticipated standard deviations for mean BMD changes at total body, spine, and trochanter (3.5, 4.5, and 6.5, respectively). The analysis yielded 98% power to detect meaningful differences in spine BMD (the primary variable) as well as at some of the secondary sites, using a 0.05 significance level (2-tailed) and a 40% dropout rate. The standard deviations were obtained from previous studies with alendronate. As it turned out, the total patient enrollment (437) was substantially greater than anticipated.

B.2.7 Changes in conduct of study/analyses

The only significant change in the conduct of this trial was the sponsor's decision to terminate the study, due to the unexpected finding of osteosarcoma in rats treated with LY333334. This has been described in detail in the sections above. The termination procedures, together with the gathering of efficacy and safety data, were essentially the same as for GHAC. In December 1998, the sponsor terminated GHAJ and instructed the investigators to have all patients complete the Early Discontinuation visit. Early Discontinuation visits that occurred after December 17, 1998 are referred to as the study closeout visits. All patients who were enrolled as of this date had completed 9 to 15 months of double-blind treatment. Approximately 90% of these patients completed study closeout visits by February 1, 1999 (within 6 weeks of discontinuation of drug), and the last patient visit occurred on March 19, 1999.

One hundred eighteen patients received the study drug following the scheduled 12-month visit, and only 1 patient had more than 14 months of exposure. This variability in exposure necessitated changes in the efficacy analyses. Thus, the sponsor added by-visit analyses for all the efficacy variables, supplementing the original analyses. Data from the study closeout visit were analyzed as well. These analyses were specified while the study was still blinded.

The "12-Month visit" (Visit 6) was conducted either as part of the regular testing schedule or as the study closeout visit. Only a small proportion of the patients had a study closeout visit at Visit 7; the sponsor gives this as the reason for the decision to use only the results from the Visits 1 through 6 (Baseline to Month 12) for analysis of the primary endpoint, although many of the tables and graphs in the submission present the results from all visits. Thus, at some point in the modification of the data analysis plan, the Month 12 data, using LOCF, became the basis for evaluation of efficacy endpoints. Final data from a number of patients whose closeout was at Visit 7 were excluded. Comments on this are provided below, as well as in the Statistics Review.

For most patients, several weeks elapsed between last dose of drug and study closeout visit. It was considered unlikely that this would affect the BMD measurements; however, alterations in bone marker responses could be expected following withdrawal of the drug, with a tendency towards underestimation of the drug effect.

Comments: There is some confusion, throughout the GHAJ report, regarding the data set that is to be used for the BMD efficacy analyses: Visit 6 with LOCF or Visit 7 with LOCF. The sponsor decided on Visit 6 for reasons given above. To the best of my knowledge, the decision to use this set of data was made without pre-submission of a formal, modified data analysis plan that was approved by our Biometrics Division. I see no *a priori* reason for choosing the Visit 6 data rather than the Visit 7 data, since the latter would truly represent the last observation for everyone. The Visit 6 results are derived from a combination of LOCF and next-to-LOCF data. The sponsor's statement that a Visit 7 analysis was precluded because only a small proportion of patients had this visit is unconvincing. It should also be noted that the subset of patients who had a Visit 7 had a longer exposure time on drug. As will become evident, the choice of data set had no meaningful effect on spine BMD efficacy outcomes, but determined the presence or absence of statistical significance for percent change at the whole body and multiple sites at the hip.

Fractures were recorded as safety outcomes. The sponsor performed by-visit analyses as well as analyses of data obtained through the study closeout visit, for non-vertebral fractures, laboratory safety assessments, vital signs and all adverse events.

B.2.8 Results

B.2.8.1 Populations enrolled/analyzed

According to the sponsor, 959 patients were "entered into the screening phase." Of these, 437 were randomized to one of 3 treatment arms of the study. Patients were treated at 37 study sites in 11 countries (number of patients entered at each site provided in NDA). Out of the 437 patients, 147 received placebo, 151 received LY333334 20 μg/day, and 139 received LY333334 40 μg/day. Disposition of patients into the 3 treatment arms of each of the 2 sub-studies is also provided (the numbers were about half for each cell).

Comments: There is no description of methodology for initial patient contact or for reasons for exclusion of the 522 individuals. Further comments on the recurrence of this omission in this and other NDAs appear above.

B.2.8.2 Discontinuations

No patient completed the intended 24 months of treatment, due to the sponsor's decision to terminate the study early. This was by far the most common reason for discontinuation (as was true of GHAC). Three hundred fifty six (81.5%) patients [130 (88.4%) in the placebo group, 123 (81.5%) in the 20μg group, and 103 (74.1%) in the 40 μg group (χ^2 p=0.008 for 40 μg vs placebo)] discontinued from the study early due to sponsor's decision. The increase in discontinuation in the 40 μg group was due to adverse events and patient decision.

The reasons for study discontinuation in substudy I, substudy II and overall are summarized in Tables GHAJ.10.2 through GHAJ.10.4. The listing of primary reasons for study discontinuation for randomized patients is provided in the NDA Appendix 16.2.2.

The reasons for study discontinuation, by group, for the entire study cohort are provided in the following table:

BEST POSSIBLE COPY

Reasons Discontinued	PLACEBO (N=147) n (%)	PTH20 (N=151) n (%)	PTH40 (N=139) n (%)	TOTAL (N=437) n (%)	P-VALUES P_PCHI	P_MHCHI	_EXACT_
Sponsor's decision	130(88.4)	123(81.5)	103(74.1)	356(81.5)	0.006	0.001	0.008
Adverse event	7(4.8)	14(9.3)	18(12.9)	39(8.9)	0.052	0.015	0.045
Patient decision	4(2.7)	8(5.3)	13(9.4)	25(5.7)	0.052	0.016	0.055
Exclusion medication	1(0.7)	1(0.7)	1(0.7)	3(0.7)	0.996	0.969	1.000
Other clinically significant							
Lab values	0(0.0)	2(1.3)	1(0.7)	3(0.7)	0.383	0.450	0.539
Death	0(0.0)	2(1.3)	0(0.0)	2(0.5)	0.149	0.974	0.333
Lack of efficacy,							
progressive disease	2(1.4)	0(0.0)	0(0.0)	2(0.5)	0.138	0.086	0.213
Unable to contact							
patient (lost to follow-up)	1(0.7)	0(0.0)	1(0.7)	2(0.5)	0.589	0.974	0.547
Noncompliance	0(0.0)	0(0.0)	1(0.7)	1(0.2)	0.342	0.208	0.318
Other	0(0.0)	0(0.0)	1(0.7)	1(0.2)	0.342	0.208	0.318
Patient moved	1(0.7)	0(0.0)	0(0.0)	1(0.2)	0.373	0.235	0.454
Physician decision	1(0.7)	0(0.0)	0(0.0)	1(0.2)	0.373	0.235	0.454
Protocol entry criteria							
not met	0(0.0)	1(0.7)	0(0.0)	1(0.2)	0.387	0.982	1.000

$ P-VALUES ARE FROM THE FOLLOWING TESTS:
P_PCHI: PEARSON'S CHI SQUARE TEST
P_MHCHI: MANTEL-HAENSZEL TEST FOR CORRELATION WITH DOSE
EXACT: FISHER'S EXACT TEST

Note the greater proportion of adverse events and discontinuations due to patient decision in the 40 µg group, relative to placebo. These seemed to increase in a dose-dependent manner, but the differences between placebo and 20 µg were not statistically significant. Two deaths were reported in the LY333334 20 µg group, and 1 in the 40 µg group. There were no deaths in the placebo group. According to the investigators, none of the 3 deaths was related to study drug. An evaluation of all adverse events and deaths is provided in the Integrated Safety Review. From the standpoint of efficacy, however, there was no imbalance in discontinuations among treatment groups that was of sufficient magnitude to affect or invalidate the results. This applies particularly to the group receiving 20 µg/day, the indicated dose of LY333334. The overall retention rate, until the sponsor discontinued the trial, was over 80% for placebo and the 20 µg group (74% for the 40 µg group).

According to the protocol, increases in serum or urinary calcium values that were above the normal range may have resulted in reductions or discontinuations in calcium supplementation or in dose of study drug (by half). The same provision applied to patients who experienced side effects likely attributable to the study drug, such as nausea, dizziness, or other expected adverse events. Of the 437 randomized patients, 37 (8.5%) had the dose of study drug permanently reduced to 50%: three patients in the placebo group, 11 in the 20 µg group and 23 in the 40 µg group. These percentages were statistically significant and dose related ($p<0.05$ for all pairwise comparisons). All of these patients discontinued the study, nearly all due to sponsor's decision.

B.2.8.3 Protocol violations

Clinically relevant protocol violations included patients who did not meet enrollment criteria, lab samples or visits either not performed or not performed within specified times, specific protocol procedures not performed, and drug accountability issues. Each of these violations is listed in the NDA in Table GHAJ.10.6. In this table, the specific violation is given, along with the subject number and treatment group. A review of these violations shows that most are minor and they appear to be evenly distributed among the treatment groups.

There were very few drug accountability violations and all appeared to involve inability to return all bottles of calcium and vitamin D. In my opinion, there is no reason to believe that protocol violations could have affected the conduct or results of this study. In addition for efficacy analyses, all data collected prior to randomization, whether or not within the specified time windows, were used as baseline data in the ITT approach.

The sponsor provides a detailed account of patient compliance with study drug. On average the patients in each treatment group took at least 79% of study drug, with no significant between-group differences found. Further details are provided in the submission.

B.2.8.4 Baseline characteristics of enrolled population

A total of 437 patients were randomized into the study. These contributed all the data for the efficacy analyses, based on ITT approach.

Data from 251 of these were used for the population pk analyses of LY333334 in osteoporotic men. All 251 patients were taking LY333334, 20 or 40 µg/day. The 251 patients ranged in age from 31 to 84 years and weighed between 47.6 and 128.9 kg.

No pharmacodynamic data sets were obtained for this study.

Demographics and baseline characteristics of the 437 randomized patients are given in the following table:

Characteristic	Substudy I (N=223)	Substudy II (N=214)	Overall (N=437)	p-value
Age (Years) (Mean ± SD)	58.94±12.69	58.42±13.29	58.68±12.98	0.917
Origin n (%)				0.718
Caucasian	222 (99.6)	211 (98.6)	433 (99.1)	
Asian	0	2 (0.9)	2 (0.5)	
Other	1 (0.4)	1 (0.5)	2 (0.5)	
Body-mass index (kg/m^2) (Mean ± SD)*	24.89±3.39	25.42±3.87	25.15±3.64	0.314
Current smoker n (% yes)	76 (34.1)	54 (25.2)	130 (29.7)	0.129
Alcohol n (% yes)	156 (70.0)	150 (70.1)	306 (70.0)	>0.99
Previous osteoporosis drug user n (% yes)	31 (13.9)	33 (15.4)	64 (14.6)	0.653
Osteoporosis type n (%)				0.999
Idiopathic	114 (51.1)	109 (50.9)	223 (51.0)	
Hypogonadal	109 (48.9)	105 (49.1)	214 (49.0)	
Previous nonvertebral fracture n (% yes)	139 (62.3)	119 (55.6)	258 (59.0)	0.529
Baseline vertebral BMD (g/cm^2)(Mean ± SD)	0.87±0.14	0.87±0.15	0.87±0.14	0.973
Dietary Calcium (g/d) (Mean ± SD)	0.81±0.55	0.86±0.52	0.84±0.54	0.693
Caffeine n (% yes)	205 (91.9)	179 (83.6)	384 (87.9)	0.112

There were no significant differences according to sub-study. Nearly all patients were Caucasian. About half the patients were hypogonadal. More than half had suffered a previous non-vertebral fracture. The mean baseline vertebral BMD in all 3 groups (about0.87g/cm2) would correspond to a T-score of –2 using Hologic data.

The sponsor provides baseline data for patients in each of the 2 sub-studies, by treatment group. There were no differences across treatment groups or between the 2 sub-studies in any category. Data for the GHAJ study group as a whole are provided in the following table:

Characteristic	Placebo (N=147)	PTH20 (N=151)	PTH40 (N=139)	Total (N=437)	p-value
Age (Years) (Mean ± SD)	58.65±12.87	59.29±13.40	58.06±12.68	58.68±12.98	0.724
Origin n (%)					0.725
Caucasian	147 (100)	149 (98.7)	137 (98.6)	433 (99.1)	
Asian	0	1 (0.7)	1 (0.7)	2 (0.5)	
Other	0	1 (0.7)	1 (0.7)	2 (0.5)	
Body-mass index (kg/m²) (Mean ± SD)*	25.21±3.61	25.37±3.72	24.86±3.60	25.15±3.64	0.483
Height cm (Mean ± SD)**	173.63±7.40	173.72±7.34	172.99±7.45	173.46±7.39	0.665
Weight kg (Mean ± SD)	75.98±11.54	76.59±12.25	74.47±12.16	75.71±11.99	0.305
Current smoker n (% Yes)	47 (32.0)	45 (29.8)	38 (27.3)	130 (29.7)	0.693
Alcohol n (% yes)	102 (69.4)	114 (75.5)	90 (64.7)	306 (70.0)	0.134
Previous osteoporosis drug user n (% yes)	17 (11.6)	22 (14.6)	25 (18.0)	64 (14.6)	0.308
Osteoporosis type n (%)					0.974
Idiopathic	74 (50.3)	78 (51.7)	71 (51.1)	223 (51.0)	
Hypogonadal	73 (49.7)	73 (48.3)	68 (48.9)	214 (49.0)	
Previous non-vertebral fracture n (% yes)	79 (53.7)	100 (66.2)	79 (56.8)	258 (59.0)	0.139
Baseline vertebral BMD (Mean ± SD)	0.85±0.14	0.89±0.15	0.87±0.14	0.87±0.14	0.053
Dietary Calcium (Mean ± SD)	0.86±0.57	0.84±0.54	0.80±0.50	0.84±0.54	0.667
Caffeine n (% yes)	130 (88.4)	128 (84.8)	126 (90.6)	384 (87.9)	0.425

There were no meaningful differences in any of these baseline characteristics, across treatment groups.

B.2.8.5 Efficacy outcomes

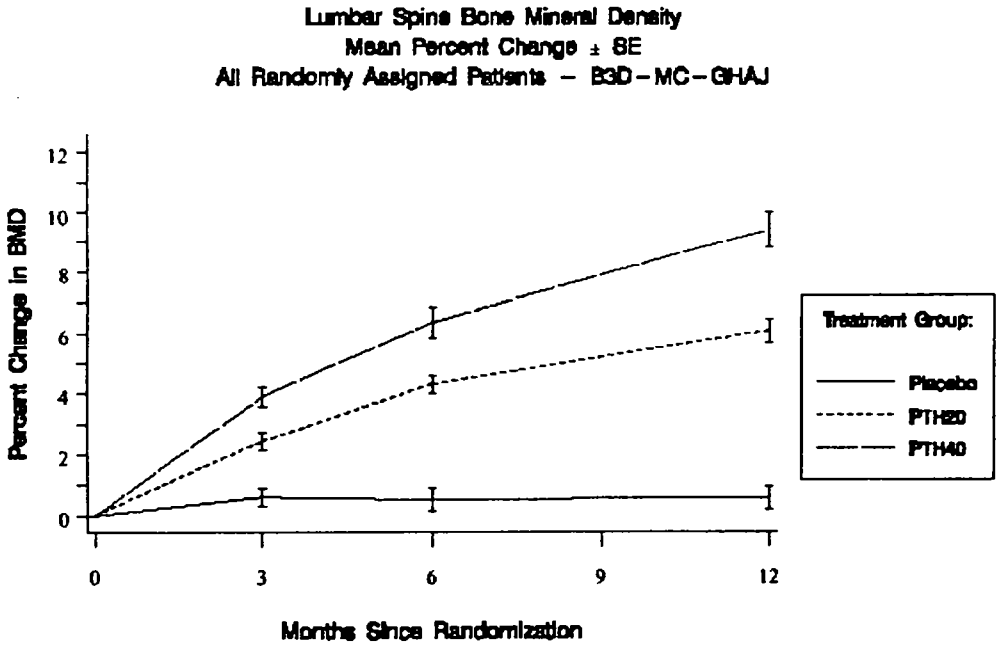
B.2.8.5.1 Bone Mineral Density

As specified in the protocol, the primary analysis of BMD data was the mean percent change from baseline to endpoint using ITT. The sponsor used data from all randomized patients with a baseline value and at least one post-baseline value up to 12 months. The patient's last post-randomization value was carried forward (LOCF). This constituted the main analysis for the sponsor's BMD efficacy claims. In addition, Visit 7 data, with LOCF, are provided (**see discussion of this issue above**). If more than one BMD assessment was made at any skeletal site at any visit, the average value was used. In addition, the BMD changes from baseline to each visit were analyzed using only patients who had data at that visit.

The analyses that the sponsor presents in this section include data from every visit, including those that were derived from the study closeout visits. Data are presented in detail by visit, for each sub-study and overall. There were no meaningful differences in outcomes between the two sub-studies; thus this review will summarize the overall results.

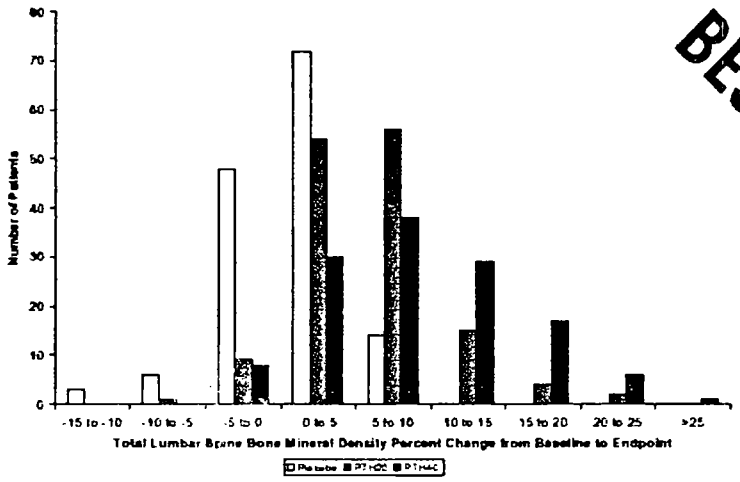
The primary efficacy outcome of GHAJ was clearly achieved; the study demonstrated a substantial and statistically significant effect of LY333334 treatment on spinal BMD. Patients treated with LY333334 20 µg/day and 40 µg/day had statistically significant, placebo-subtracted increases in lumbar spine BMD of 5.19% and 8.21%, respectively, at the 12-month endpoint ($p<0.001$ for each comparison with placebo, and $P<0.001$ for comparison between LY333334 doses). The baseline-to-endpoint (i.e., including Visit 7 data) results for spine BMD were essentially the same, with placebo-subtracted differences of 5.35% for the 20 µg group and 8.51% for the 40 µg group ($p<0.001$ for both comparisons with placebo).

The time course for BMD changes in the LY333334 and placebo groups are shown in the next figure. At each time point after baseline, the comparisons between the LY333334 groups and between each active group and placebo were statistically significant ($p<0.001$).



Responder analysis for lumbar spine BMD responses: Approximately 39.9% of patients in the placebo group had a decrease in lumbar spine BMD at study endpoint. In contrast, a decrease in vertebral BMD was seen in 7.1% or 6.2% of the patients treated with either 20 μg or 40 μg of LY333334, respectively. In 9.8% of patients in the placebo group, lumbar spine BMD increased by ≥ 5%, compared with 54.6% of patients receiving LY333334 20 μg and 70.5% of those receiving 40 μg. In the 40 μg group, over 40% had BMD increases that were in excess of 10% over baseline (as opposed to about 15% of the 20 μg group with BMD increases of over 10 %).

The data describing the numbers of patients with a given BMD response, by treatment group, are presented in the next figure:

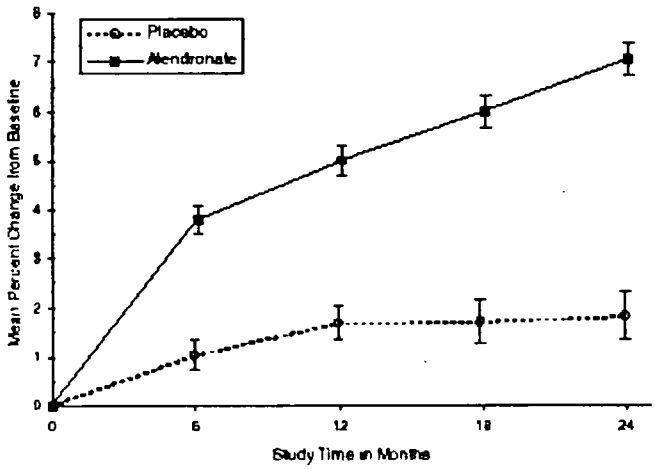


As described below, there were no statistically significant differences between the hypogonadal and idiopathic osteoporotic subgroups in lumbar spine BMD responses.

Comments: These BMD responses can be compared with results from the male osteoporosis trial of alendronate, the only drug that is currently approved for this indication. In that trial, which studied a comparable group of patients over 24 months, the mean placebo-subtracted increase in lumbar spine BMD was 5.32% (at 24 months), with increases of about 4.5% at 12 months. A responder analysis of the 24-month data showed that 87% of alendronate-treated patients increased BMD by 3% or more (vs 29% of placebo patients). About 70% of patients had BMD increases of 5% or more, vs about 17% of placebo patients. Thus, for the primary endpoint, GHAJ showed that LY333334, 20 μg/day for up to approximately 12 months,

achieved BMD increases that were similar to those produced by alendronate 10 mg/day for 24 months. Similar to the alendronate trial, the BMD responses of the hypogonadal group were essentially the same as those of the eugonadal patients.

The lumbar spine BMD responses to alendronate, 10 mg/day, and placebo, In men with osteoporosis are shown in the following figure, taken from Merck's sNDA 20560-023:



APPEARS THIS WAY
ON ORIGINAL

For the secondary BMD endpoints hip and total body, statistically significant % increases from baseline, relative to placebo, were found for the 20 µg and 40 µg groups, using the Month 12 data with LOCF. At the hip, the placebo-subtracted increases were 0.73% for the 20 µg group and 1.92% for the 40 µg group. For the total body, the placebo-subtracted increases were 0.83% for the 20 µg group and 0.87% for the 40 µg group (the placebo group lost 0.33%, whereas both LY333334 groups gained about 0.5% in total body BMD). For the secondary endpoint BMD at the radius, the % differences from baseline did not differ between either LY333334 treatment group and placebo (both for the distal and

ultra-distal radius). Of note, the placebo and 20 g groups lost BMD at both these sites. The 40 µg group lost BMD at the distal, but not ultradistal, radius. Data are summarized in sponsor's and reviewer's tables below.

The difference between the two LY333334 doses, in % change from baseline BMD, was statistically significant at the total hip (p=0.011), but not for the total body or radius.

BMD was measured at several other skeletal sites, including femoral neck, trochanter, intertrochanter, and Ward's triangle. Using the 12-month data with LOCF, there were statistically significant differences between each treatment group and placebo at the femoral neck (1.08% and 2.49% for the 20 µg and 40 µg groups, respectively) and between the two treatment groups. However, at the other three sites (all at the hip), the differences between the 20 µg group and placebo were not statistically significant (data presented in tables below). Using endpoint data (with LOCF), statistical significance differences from PBO were found at the femoral neck only.

Comments: These four skeletal sites were not defined as specific endpoints or objectives in the protocol, and the sponsor's use of p-values for BMD analyses at these sites is not appropriate (the use of p-values without adjustment for multiplicity, even for defined secondary endpoints, is also not justified). However, the methodology section of the protocol did specify that BMD was to be measured at these skeletal sites; indeed, some of the measurements were made in a pre-specified sub-set of patients.

A recurrent problem with this application is the confounding of specific objectives on the one hand, with a multiplicity of measured outcomes, on the other (e.g., my earlier comments in the review of GHAC). In this instance, the specific secondary objectives, as stated in the GHAJ protocol, include measurements of "hip BMD," without specifying which of the five hip measurements would be used in the determination of efficacy (e.g., Main Study Report, Objectives, page 73; *ibid.*, Investigational Plan, page 75). During the study, the sponsor measured BMD at the total hip, femoral neck, Ward's triangle, trochanter, and intertrochanter. Statistically significant % increases in BMD, relative to placebo, were found at the total hip and femoral neck (or femoral neck only, depending on data set used), but not at the trochanter, intertrochanter, or Ward's triangle. These negative results do not appear in the sponsor's efficacy claims for LY333334, as shown in the following table, reproduced from the proposed label.

	FORTEO N=151	Placebo N=147
Lumbar Spine BMD	5.7[a]	0.5
Femoral Neck BMD	1.4[b]	0.4
Total Hip BMD	1.1[b]	0.4

[a] $p<0.001$ compared with placebo

[b] $p<0.05$ compared with placebo

For the 3 skeletal sites that are not mentioned in the sponsor's claims (trochanter, intertrochanter, and Ward's triangle), there were numerical increases in the 20 μg group, relative to placebo, with further increases that attained statistical significance in the 40 μg group (all data are shown in the sponsor's tables below). At two of the three sites (trochanter and Ward's triangle) comparisons of absolute BMD changes (in g/cm^2) reached statistical significance.

The point is not whether the femoral neck is more important clinically than Ward's triangle or the trochanter, but whether consistent and proper analytical procedures have been maintained. The ideal approach to reporting secondary outcomes in osteoporosis trials has not been established. One can debate the proper use of p-values for evaluation of the significance of multiple secondary endpoints. In my opinion, a correction for multiple comparisons should be employed, unless there are ancillary supportive data. However, the use of p-values to describe endpoints that have not been explicitly pre-specified as outcomes is unacceptable. Furthermore, the practice of choosing favorable results from among a multiplicity of outcomes, while neglecting to mention some of the less favorable outcomes, is misleading.

As commented in the GHAC review, at no point is a hypothesis clearly stated. A hypothesis should not be regarded simply as a re-wording of an efficacy objective. The presentation of a hypothesis for each endpoint adds rigor and credibility to the data analysis and permits the use of p-values (or comparisons using confidence intervals), when appropriate and pre-specified. The lack of a hypothesis facilitates both the introduction of endpoints that are not pre-specified and the suppression of results that are unfavorable from the sponsor's standpoint.

Perhaps one approach to this overall problem would be to list all pre-specified secondary efficacy results, without the use of p-values.

To summarize the % increases in BMD at each skeletal site, by treatment group at the 12-month endpoint, I have reproduced the sponsor's table, below:

Variable	Placebo (N=147)	PTH20 (N=151)	PTH40 (N=139)	P-Value (Treatment Comparison)			
				Overall	Placebo vs PTH20	Placebo vs PTH40	PTH20 vs PTH40
Lumbar Spine (L-1 through L-4)							
n	143	141	129	—	—	—	—
Mean baseline (g/cm²)	0.85±0.14	0.89±0.15	0.87±0.14	0.016	0.005	NS	NS
Mean change (g/cm²)	0.01±0.03	0.05±0.04	0.07±0.05	<0.001	<0.001	<0.001	<0.001
Mean percent change	0.54±4.19	5.73±4.46	8.75±6.25	<0.001	<0.001	<0.001	<0.001
Total Hip							
n	137	135	125				
Mean baseline (g/cm²)	0.83±0.11	0.84±0.10	0.83±0.11	NS	NS	NS	NS
Mean change (g/cm²)	0.00±0.02	0.01±0.02	0.02±0.03	<0.001	0.017	<0.001	0.017
Mean percent change	0.41±2.77	1.14±2.89	2.33±4.51	<0.001	0.040	<0.001	0.011
Femoral Neck							
n	137	135	125	—	—	—	—
Mean baseline (g/cm²)	0.70±0.11	0.71±0.10	0.70±0.11	NS	NS	NS	NS
Mean change (g/cm²)	0.00±0.03	0.01±0.03	0.02±0.04	<0.001	0.013	<0.001	0.032
Mean percent change	0.36±3.95	1.44±3.61	2.85±6.07	<0.001	0.038	<0.001	0.016
Trochanter							
n	137	135	125	—	—	—	—
Mean baseline (g/cm²)	0.65±0.11	0.66±0.10	0.65±0.12	NS	NS	NS	NS
Mean change (g/cm²)	0.01±0.02	0.01±0.03	0.01±0.03	NS	NS	0.024	NS
Mean percent change	0.95±3.40	1.25±4.15	1.98±5.16	NS	NS	0.044	NS

Variable	Placebo (N=147)	PTH20 (N=151)	PTH40 (N=139)	P-Value (Treatment Comparison)			
				Overall	Placebo vs PTH20	Placebo vs PTH40	PTH20 vs PTH40
Intertrochanter							
n	137	135	125	—	—	—	—
Mean baseline (g/cm²)	0.96±0.13	0.98±0.13	0.97±0.14	NS	NS	NS	NS
Mean change (g/cm²)	0.00±0.03	0.01±0.03	0.02±0.04	<0.001	0.030	<0.001	0.041
Mean percent change	0.48±2.91	1.20±3.07	2.32±4.57	<0.001	NS	<0.001	0.024
Ward's Triangle							
n	137	135	125				
Mean baseline (g/cm²)	0.51±0.12	0.51±0.11	0.50±0.13	NS	NS	NS	NS
Mean change (g/cm²)	0.00±0.04	0.01±0.04	0.03±0.05	<0.001	0.044	<0.001	0.003
Mean percent change	0.21±8.64	2.48±7.20	6.19±10.21	<0.001	NS	<0.001	0.001
Whole body[a]							
n	87	84	83	—	—	—	—
Mean baseline (g/cm²)	1.07±0.09	1.08±0.09	1.07±0.08	NS	NS	NS	NS
Mean change (g/cm²)	0.00±0.03	0.01±0.03	0.01±0.03	0.025	0.026	0.015	NS
Mean percent change	0.33±2.51	0.50±2.99	0.54±2.45	0.039	0.039	0.021	NS
Ultradistal Radius (Forearm)[a]							
n	93	89	85	—	—	—	—
Mean baseline (g/cm²)	0.43±0.06	0.44±0.07	0.43±0.06	NS	NS	NS	NS
Mean change (g/cm²)	0.00±0.01	-0.00±0.01	0.00±0.02	NS	NS	NS	NS
Mean percent change	0.53±3.28	-0.40±3.15	0.54±5.98	NS	NS	NS	NS

Variable	Placebo (N=147)	PTH20 (N=151)	PTH40 (N=139)	P-Value (Treatment Comparison)			
				Overall	Placebo vs PTH20	Placebo vs PTH40	PTH20 vs PTH40
Distal Radius (Forearm)[a]							
n	93	89	85	—	—	—	—
Mean baseline (g/cm²)	0.78±0.12	0.78±0.12	0.77±0.11	NS	NS	NS	NS
Mean change (g/cm²)	-0.00±0.02	-0.00±0.02	-0.01±0.02	NS	NS	NS	NS
Mean percent change	-0.18±2.08	-0.47±2.21	-0.67±2.56	NS	NS	NS	NS

Comments: Consistent with the results of GHAC and the phase 2 clinical pharmacology data, the BMD increases associated with LY333334, 40 µg were substantially and statistically significantly greater than with 20 µg at all sites except for the total body and the radius. Sites that are rich in cancellous bone

respond particularly well to LY333334, in a dose-dependent manner within this dose range. It is likely that the substantial increase in efficacy of the 40 μg dose, relative to 20 μg, is partially due to the somewhat lower systemic exposure to LY333334 in men (see above population pk results).

In the following (reviewer's) table, I have summarized the placebo-subtracted 12-month endpoint BMD changes at each skeletal site, for the placebo and 20 μg groups:

SKELETAL SITE	MEAN (PLACEBO-SUBTRACTED) % BMD INCREASE	TREATMENT COMPARISON (20 μg vs PLACEBO)
LUMBAR SPINE	5.19	p<0.001
TOTAL HIP	0.73	p<0.040
FEMORAL NECK	1.08	p<0.038
TROCHANTER	0.30	NS
INTERTROCHANTER	0.72	NS
WARD'S TRIANGLE	1.77	NS
WHOLE BODY	0.83	p<0.039
ULTRADISTAL RADIUS	0.13	NS
DISTAL RADIUS	-.029	NS

Comments: The sponsor also analyzed the results using the Visit 7 endpoint, using LOCF. Using these data, for the 20 μg group, statistical significance (compared to placebo) was lost at the total hip and whole body (% change but not absolute change in BMD). For the 40 μg group, the statistically significant comparisons that were seen at Month 12 were maintained at Visit 7, with the exception of the trochanter (% change but not absolute change in BMD). I have summarized the placebo-subtracted BMD changes at endpoint (Visit 7 with LOCF) for the 20 μg group in the following (reviewer's) table:

SKELETAL SITE	MEAN (PLACEBO-SUBTRACTED) % BMD INCREASE	TREATMENT COMPARISON (20 μg vs PLACEBO)
LUMBAR SPINE	5.35	p<0.001
TOTAL HIP	0.63	NS
FEMORAL NECK	1.24	p<0.029
TROCHANTER	0.24	NS
INTERTROCHANTER	0.57	NS
WARD'S TRIANGLE	1.76	NS
WHOLE BODY	0.76	NS
ULTRADISTAL RADIUS	-0.19	NS
DISTAL RADIUS	-.031	NS

In my opinion, these are the BMD results that should be used in evaluation of LY333334 efficacy.

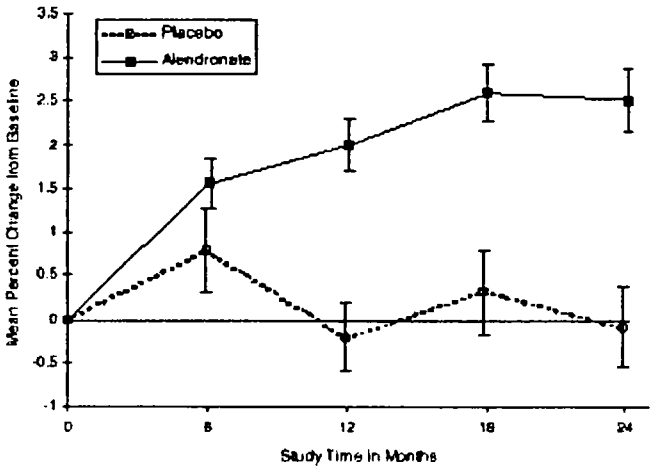
Much larger, and statistically significant (relative to placebo), increases in BMD were found in the 40 μg group at all the above skeletal sites except for the trochanter (p=0.06), ultradistal, and distal radius at the study endpoint. The results are similar to those presented at Month 12 (see sponsor's table above).

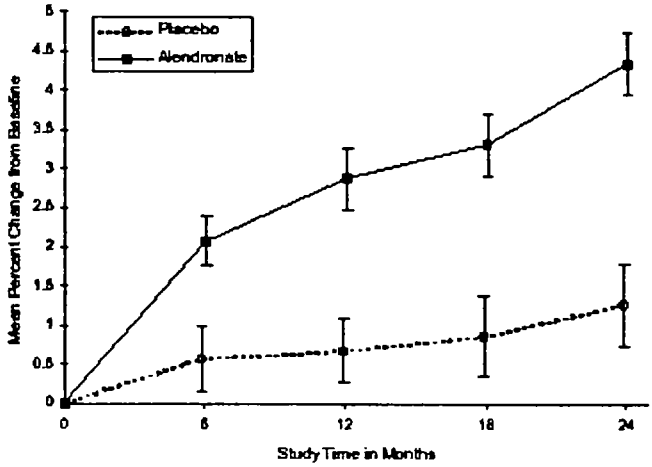
It is of interest to compare these results to the outcomes of the alendronate male osteoporosis trial (sNDA 20560-023). The following table is taken from my review of that submission. The numbers represent % increase in BMD from baseline to study end at 24 months. The placebo-subtracted difference for the lumbar spine is 5.25%.

SKELETAL SITE	ALN (% BMD change from baseline)	PBO (%BMD change from baseline)
Lumbar spine	7.07 a, b	1.82 a
Femoral neck	2.52 a, b	- 0.08
Trochanter	4.35 a, b	1.27 a
Total hip	3.12 a, b	0.58
Ward's triangle	4.14 a	2.86 a
Total body	1.95 a, b	0.39

a= statistically significantly different from baseline
b= statistically significantly different from PBO
ALN=alendronate, 10 mg/day; PBO=placebo

To give some idea of the time course of changes at two of the skeletal sites in the alendronate study, the following figures show BMD changes at the femoral neck (left panel) and trochanter (right panel) in the alendronate male osteoporosis trial (figures taken from sNDA 20560-023):





Comments: Judging from these comparative data, alendronate, 10 mg/day, seems to be a little less effective than LY333334 20 µg/day at the lumbar spine and probably more effective at other skeletal sites. The data are presented to place the LY333334 results in some perspective. We have no data comparing efficacy results of these doses of the two agents in the same trial in men.

Results of changes in bone mineral content, by skeletal site and treatment group, are also presented. Again, there were substantial and statistically significant increases in BMC at the lumbar spine in both treatment groups, compared to placebo. For the 20 µg group, comparisons with placebo lost statistical significance at all other sites. A review of these data suggests that the loss of significance was partly due to increases in variability. There was also a sizeable increase in placebo BMC in at least one site and diminished efficacy in the treatment groups at another site. The overall numerical trends appeared to follow the BMD data at all sites except for trochanter and Ward's triangle.

Changes in bone mineral area were also analyzed (the DEXA measures bone mineral content and divides this by the measured bone mineral area, yielding BMD). Increases in BMA in one treatment group could cause some underestimation of the treatment effects on BMD. There were significant increases in BMA at the lumbar spine in the two LY333334 treatment groups, compared to placebo ($p<0.001$). There were no significant differences between the two LY333334 groups. There were no meaningful changes at other skeletal sites. It is therefore possible that the BMD effects of LY333334 treatment represent an underestimation of the change in the lumbar spine mineral space (i.e., size of the vertebral space and content of mineral in that space). A

comparison of changes in BMD and BMC at Month 12 (with LOCF), by treatment group, is provided in the next table:

Site	Mean Percent Change from Baseline to Endpoint						P-Value (Pairwise Comparison)					
	Placebo		PTH20 group		PTH40 group		Placebo vs PTH20		Placebo vs PTH40		PTH20 vs PTH40	
	BMD	BMC	BMD	BMC	BMD	BMC	BMD	BMC	BMD	BMC	BMD	BMC
Lumbar Spine	0.54	0.67	5.73	7.52	8.75	11.10	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Total Hip	0.41	1.34	1.14	1.95	2.33	3.00	0.040	NS	<0.001	0.018	0.011	NS
Femoral Neck	0.36	0.11	1.44	1.39	2.85	3.29	0.038	NS	<0.001	<0.001	0.016	0.019
Trochanter	0.95	4.54	1.25	3.02	1.98	2.17	ONS	ONS	ONS	ONS	ONS	ONS
Intertrochanter	0.48	0.50	1.20	1.48	2.32	3.05	NS	NS	<0.001	<0.001	0.024	0.048
Ward's Triangle	0.71	0.36	2.48	0.34	6.19	0.41	NS	ONS	<0.001	ONS	0.001	ONS
Whole body[a]	-0.33	-0.31	0.50	0.62	0.54	1.00	0.039	NS	0.021	0.008	NS	NS
Ultradistal Radius[a]	-0.53	-0.55	-0.40	-0.68	0.54	0.08	NS	NS	NS	NS	NS	NS
Distal 1/3 Radius[a]	-0.18	0.14	-0.47	-0.43	-0.67	-0.89	ONS	NS	ONS	0.002	ONS	NS

Abbreviations: PTH20 = LY333334 20μg/day; PTH40 = LY333334 40μg/day; BMD = bone mineral density; BMC = bone mineral content; NS = not statistically significant at p=0.05; ONS = overall therapy is not significant.
[a] Whole body and radius BMC were measured in a subset of patients.

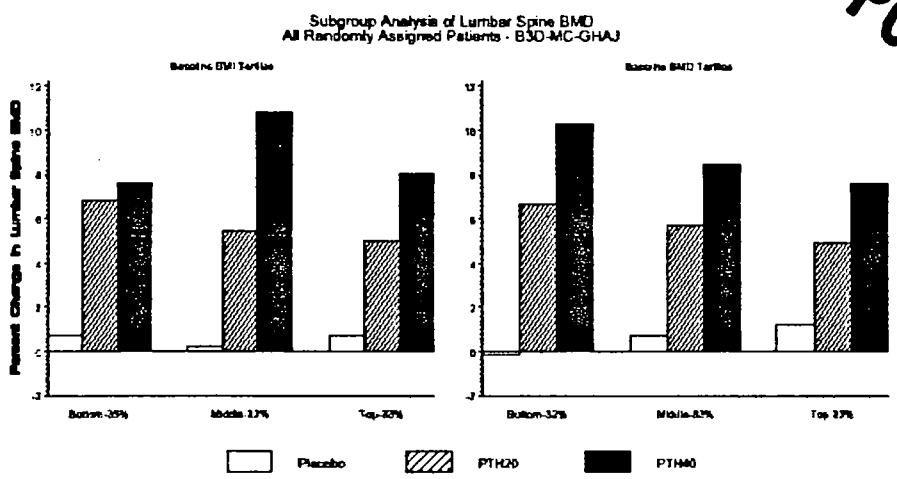
The % increases in BMC at the lumbar spine were higher than the corresponding increases in BMD. The relationship between these parameters varied among the remaining skeletal sites and according to LY333334 dose.

The sponsor presents by-visit analyses of changes in BMD at all skeletal sites. The summary data are accompanied by statistical analyses. The data do not add to the evaluation of endpoint BMD outcomes and will not be reviewed here.

Subgroup analyses:

Subgroup-by-treatment interactions were judged significant if $p<0.10$. For the six subgroups [age, BMI, baseline vertebral BMD, previous non-vertebral fracture, baseline free testosterone, and osteoporosis type (idiopathic or hypogonadal)], there was no significant interaction on BMD at any of the five skeletal sites listed (BMD efficacy at the spine, hip, femoral neck, whole body, and wrist), with two exceptions. These were interactions of spine BMD efficacy with BMI ($p=0.017$) and baseline vertebral BMD ($p=0.072$). The interaction between therapy and baseline BMD tertile was significant, but each dose of LY333334 had a statistically significant effect on spine BMD regardless of baseline tertile. Similarly, there was a significant interaction between therapy and baseline BMI tertile, but both LY333334 doses and significant therapeutic effects in patients in all three BMI tertiles.

Because BMD and fracture efficacy have occasionally varied with baseline BMD in other clinical trials, I have reproduced the sponsor's figure showing % spinal BMD increases at endpoint, by treatment group, according to baseline BMI tertile (left panel) and BMD tertile (right panel):



The sponsor also presents results obtained for subgroup analyses of BMC. None of the subgroup analyses for BMC were significant except for age for femoral neck BMC and whole body BMC, and baseline BMI and baseline vertebral BMD for spine BMC. There were interactions of baseline BMI and baseline vertebral BMD with BMC responses at the spine. However, both LY333334 doses had a significant therapeutic effect in each BMI tertile and in each baseline vertebral BMD tertile.

At the femoral neck, the youngest and oldest patients showed a significant treatment difference, but there was no significant effect of therapy among patients in the middle age tertile. However, closer examination of the data shows that, although there was an overall treatment effect in the lowest tertile, this was entirely due to the increase in BMC in the 40 µg group. There was no effect of 20 µg on femoral neck BMC, relative to placebo. For whole body BMC, the youngest and oldest (lowest and highest age tertiles) patients did not demonstrate a therapeutic effect of LY333334, whereas a beneficial effect was seen in the middle tertile. Examination of these data shows that there was an increase in both dose groups in the middle and top tertiles, with a decline in the placebo. In contrast, there was a decrease in the 40 µg group in the lowest age tertile.

Comments: The meaning of the age subgroup interactions in the femoral neck and total body BMC results is not clear. There are no obvious patterns, and there is no relationship of these findings to the subgroup analyses of BMD responses. It is possible that some or all of these statistical associations are the result of multiple comparisons.

B.2.8.5.2 Biochemical markers of bone formation and resorption

The four biochemical markers (PICP, BSAP, NTX, and DPD), plus 1,25-dihydroxyvitamin D, were measured at baseline, Months 1, 3, 6, and 12, and at Early Discontinuation or study closeout visit. As discussed above, the last visit that occurred in this study was Visit 7. Approximately 1/3 of the patients had the 12-month visit (Visit 6). For many of these patients data from Visit 6 were carried forward to Visit 7. Baseline to 12-month visit (Visit 6) changes are described in the NDA as the primary analysis, since these changes more accurately reflect the effects of therapy. In the sponsor's summary tables LOCF data are used in the analyses and endpoint changes refer to baseline to 12-month changes.

Because of the skewed nature of the distributions of biochemical marker data, the sponsor described the results in terms of median and % median changes from baseline.

Results of this analysis showed that the biomarker responses to LY333334 in osteoporotic men were similar to those in women. Consistent with the anabolic action of the drug, there were early and robust increases in the markers of bone formation, followed by increases in the two resorption markers.

In the placebo group, there were statistically significant median % decreases from baseline in BSAP at 6 and 12 months ($p<0.001$). At 12 months, the median serum BSAP level was16.3% below baseline. In contrast, both LY333334 treatment demonstrated statistically significant median % increases from baseline in serum BSAP as early as the 1-month visit, and at every scheduled visit ($p<0.001$ for all visits). At the 12-month visit, the median BSAP level for the 20 µg group was 28.8% above baseline; for the 40 µg group the median BSAP increased to 59.3% above baseline. At all visits (1-, 3-, 6-, and 12-month visits), the group receiving 20 µg had higher median percent change in BSAP levels than in placebo, and the 40 µg group had higher levels than the 20 µg group ($p<0.001$ for each comparison at all visits).

Serum procollagen I carboxy-terminal propeptide (PICP) also showed an early response to LY333334 treatment. The responses were early and tended to return towards baseline after several months of treatment. The responses of PICP to 40 µg LY333334 were greater than those that were found with 20 µg group. The 20 µg group demonstrated an initial increase that peaked at 1 month (approximately 34% over baseline, $p<0.001$). After the peak at the 1-month visit, the PICP levels declined in the 20 µg group, and at 12 months the median level was approximately 13% below baseline ($p<0.001$). In the 40 µg group, there was a significant increase in PICP, peaking at 78% at 1 month ($p<0.001$). After the peak, the levels declined, but remained statistically significantly greater than baseline at subsequent visits until Month 12, when the median PICP level was approximately 1% below baseline (NS). There were statistically significant differences between treatment groups in median percent change of PICP levels at 1, 3, 6 and 12 months ($p\leq0.001$ for all comparisons).

The prompt response of anabolic markers to LY333334 was followed by slower increases in levels of the bone resorption markers NTX and DPD. In response to both doses of LY333334, there were increases over baseline in both resorption markers, beginning at 1 month. However, for NTX, the peak responses were seen at 12 months for the 20 µg group (+57%) and at 6 months for the 40 µg group (+155%). A similar, dose- and time-dependent pattern was seen for DPD. The next table summarizes the median change and median % change from baseline to endpoint (12 months) for all four markers. With the exception of the PICP responses in the 40µg group at endpoint, there were statistically significant increases in levels of markers for both LY333334 doses, relative to baseline and relative to placebo at endpoint.

Variable	Placebo (N=147)	PTH20 (N=151)	PTH40 (N=139)	P-Value (Treatment Comparison)			
				Overall	Placebo vs PTH20	Placebo vs PTH40	PTH20 vs PTH40
Serum BSAP (µg/L)							
n	143	144	133	—	—	—	—
Median baseline	11.90 ± 0.43	10.25 ± 0.63	10.31 ± 0.87	NS	0.016	NS	NS
Median change	-1.65 ± 0.39	2.44 ± 0.41	5.00 ± 0.68	<0.001	<0.001	<0.001	<0.001
Median percent change	-16.33 ± 3.26	24.39 ± 8.06	60.54 ± 11.19	<0.001	<0.001	<0.001	<0.001
Serum PICP (ng/ml)							
n	143	144	132	—	—	—	—
Median baseline	120 ± 2.68	118.50 ± 2.32	117.50 ± 4.72	NS	NS	NS	NS
Median change	-1.00 ± 2.29	-14.00 ± 1.47	0.50 ± 3.73	0.002	0.004	NS	0.002
Median percent change	-1.14 ± 1.72	-11.71 ± 3.14	0.43 ± 3.67	0.002	0.002	NS	0.002
NTX (nmBCE/mmol creatinine)							
n	139	140	128	—	—	—	—
Median baseline	34.45 ± 1.36	33.90 ± 1.31	30.80 ± 1.02	NS	NS	NS	NS
Median change	1.70 ± 0.77	12.45 ± 1.91	28.50 ± 3.78	<0.001	<0.001	<0.001	<0.001
Median percent change	4.29 ± 2.17	47.99 ± 12.54	106.94 ± 22.98	<0.001	<0.001	<0.001	<0.001
DPD (nmol/mmol creatinine)							
n	139	140	126	—	—	—	—
Median baseline	4.35 ± 0.17	4.40 ± 0.12	4.30 ± 0.10	NS	NS	NS	NS
Median change	0.60 ± 0.09	1.65 ± 0.25	3.08 ± 0.62	<0.001	<0.001	<0.001	<0.001
Median percent change	13.56 ± 5.74	35.00 ± 3.30	72.73 ± 6.71	<0.001	<0.001	<0.001	<0.001

Abbreviations: N = number of patients, PTH20 = LY333334 20-µg/day, PTH40 = LY333334 40-µg/day;n= number of patients that have the variable assessed; BSAP = bone-specific alkaline phosphatase; NS = not significant; PICP = Serum Procollagen I C-terminal propeptide; NTX = Urinary N-telopeptide/creatinine ratio; BCE = bone collagen equivalent; DPD = Urinary free deoxypyridinoline/ creatinine ratio.

For 1,25-dihydroxyvitamin D, there were no statistically significant changes from baseline in the placebo group at any scheduled visit during the trial. For the LY333334 20 µg group, the % change from baseline was statistically significant at Months 1, 3, and 12 ($p<0.001$), with a peak increase of 22% at Month 1. For the 40 µg group, the % change from baseline was statistically significant at Months 1 through 12 ($p< 0.010$), with a peak increase of 27% at Month 1. The % change from baseline was statistically significantly greater in the 20 µg group than in placebo at Months 1, 3, and 6 ($p< 0.014$). The % change for the 40 µg group was statistically significantly greater than for placebo at Months 1 through 12 ($p< 0.046$). The changes in the 20 and 40 µg groups were not statistically significantly different at any time point.

Subgroup analyses were performed for all four biomarkers and for 1,25-dihydroxyvitamin D. A subgroup-by-treatment analysis was significant if $p<0.010$. The following table summarizes results for this analysis:

Subgroup	BSAP	PICP	NTX	DPD	1.25-di-Vit D
Age (yrs)	NS	NS	NS	NS	NS
Body mass index (kg/m^2)	NS	NS	NS	NS	NS
Vertebral BMD (g/cm^2)	NS	NS	NS	NS	NS
Previous NV fracture	0.089	NS	NS	NS	NS
Baseline free testosterone	NS	0.027	NS	NS	NS
Osteoporosis type	0.033	0.065	NS	NS	0.098

Abbreviations: BSAP = bone-specific alkaline phosphatase; PICP = Serum Procollagen I C-terminal propeptide; NTX = Urinary N-telopeptide/creatinine ratio; DPD = Urinary free deoxypyridinoline/creatinine ratio; 1.25-di-Vit D= 1,25-Dihydroxyvitamin D; NS = Nonsignificant subgroup-by-treatment interaction; BMD = bone mineral density; NV = non-vertebral.

Among the multiple comparisons, there were interactions with osteoporosis type. LY333334 increased BSAP in both eugonadal and hypogonadal patients. For PICP, LY333334 40 µg, increased PICP in the eugonadal, but not the hypogonadal, group. LY333334 20 µg did not increase PICP in either group.

B.2.8.5.3 Health-related Quality of Life (HRQOL) Indicators

A listing of the HRQOL instruments is provided in the review of GHAC, above. For the male osteoporosis study, GHAJ, the HRQOL data from the 5 instruments were essentially stable throughout the course of the study, and the sponsor observed no consistent changes or trends in any of the domains, as a result of treatment with LY333334. As noted in GHAC, two of the QOL instruments were osteoporosis-specific.

B.2.8.5.4 Population Pharmacokinetics and Pharmacokinetic Modeling

The sponsor presents a comprehensive summary of the population pk data for males with osteoporosis.

Following s.c. injection of a 20 µg dose of LY333334, the median peak serum concentration of the drug was 121.2 pg/mL. The median total systemic exposure, estimated as AUC, was 208.6 pg.hr/mL. The model suggests that in most males, serum LY333334 concentrations will be [redacted] (the lower limit of quantitation with the sponsor's assay) by 2 hours after a 20 µg dose.

The apparent volume of distribution (V/F) correlated with body weight. The predicted V/F decreased from 131 L for a 74.0 kg individual (the median value of the population) to 90 L for a patient weighing 48.2 kg, the population minimum value. However, when normalized to body weight, V/F was essentially the same across the range of weights in the population (approximately 1.8 L/kg).

The effect of body weight on V/F does not alter AUC significantly, but it may affect C_{max}. The predicted C_{max} after injection of 20 µg into the abdominal wall for a patient weighing 48.2 kg is 185.0 pg/mL; for a patient weighing 74.0 kg, the predicted peak concentration is 132.4 pg/mL. These effects of body weight on V/F, peak serum concentrations are not considered by the sponsor to be clinically significant (see concentration-time curves below).

The volume of distribution was found to be approximately 30% higher in patients who injected the dose into the thigh, compared with the abdomen. This resulted in a C_{max} that was lower after injection into the thigh. However, as shown in the combined population pharmacodynamic analyses, the BMD and biomarker responses in patients injecting into the abdomen were essentially the same as those in patients injecting into the thigh. Therefore, site of injection does not result in a clinically important effect on the disposition of LY333334.

The sponsor also studied the effect of creatinine clearance on the disposition of LY333334. In the GHAJ population, creatinine clearance, calculated from a 24-hour urine collection, ranged from 40.9 to 310.1 mL/min. The clearance of LY333334 correlated with creatinine clearance. The relationship between the two parameters is shown in the following table:

Creatinine Clearance (mL/min)		Population Estimate of LY333334 Clearance (L/hr)
40.9	(population minimum)	64.8
60.3	(5th percentile)	73.8
123.2	(median)	93.9
199.3	(95th percentile)	110.3
310.1	(population maximum)	128.0

As shown here, the predicted LY333334 CL/F decreased from 93.9 L/hr to 64.8 L/hr as creatinine clearance declined from the population median of 123.2 mL/min to the population minimum of 40.9 mL/min (i.e., a 67% decrease in CLcr resulted in a 31% decrease in clearance of LY333334).

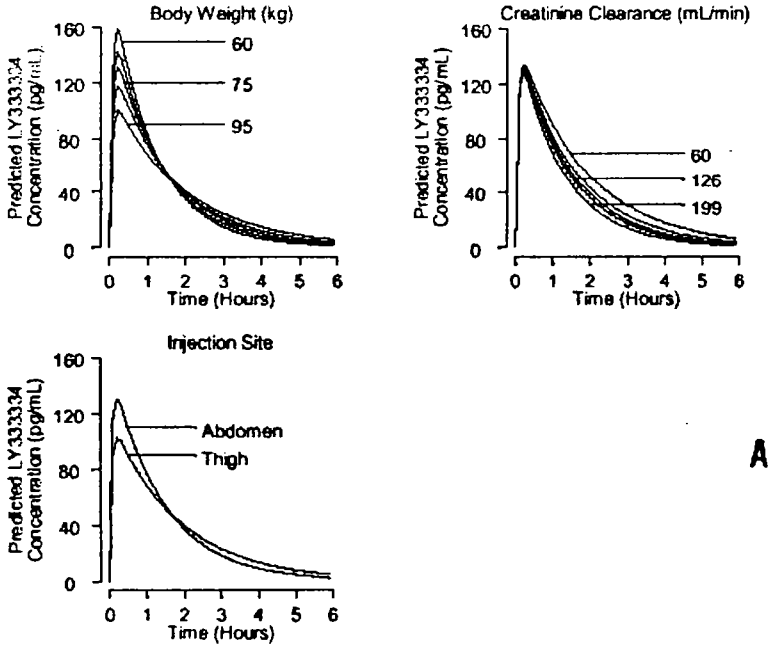
It is expected that changes in LY333334 CL/F would affect systemic exposure to the drug. For a patient with CLcr of 40.9 mL/min the predicted AUC is 309 pg•hour/mL; for a CLcr of 123.2 mL/min, the corresponding AUC is 213 pg•hour/mL. Thus a 70% reduction in renal function would be predicted to increase systemic exposure by about 45-50%.

According to the sponsor, the partial dependence of AUC on creatinine clearance over this range does not translate into clinically meaningful outcomes. For example, the addition of CLcr as a covariate in the pk model resulted in only a small reduction in inter-patient variability of LY333334 clearance. This suggests that CLcr is not an important factor in an individual's systemic exposure to

LY333334, over this range of renal function. In addition, for individuals in the lowest 5th percentile of CLcr values, there were no episodes of hypercalcemia or serious adverse events that were considered to be drug-related.

Comments: These predictions are reasonable. To give some indication of the effects of body weight, creatinine clearance, and injection site on the clearance of LY333334, I have reproduced the sponsor's concentration-time curves below. In the upper two panels, the curves represent the mean, 5th, 25th, 50th, and 95th percentile values for the two covariates. The numbers designate the 5th, mean, and 95th percentile values. The predicted results are for a 20 μg dose of LY333334. Except where noted, the results are for a patient weighing 75 kg, with a creatinine clearance of 126 ml/min, and injecting the drug s.c. into an abdominal site.

Note that the assay employed in the pk studies (from which the actual numbers were derived) could not detect concentrations of LY333334 that were < 50 pg/ml. Adjusting for molecular weight, the upper limit of normal for endogenous PTH 1-84 (65 pg/ml) would correspond to a drug concentration of 26 pg/ml. It is debatable, however, whether this consideration of assay sensitivity would affect the final pharmacokinetic models.



APPEARS THIS WAY
ON ORIGINAL

The sponsor also attempted to evaluate the effects of several other parameters on LY333334 disposition. The effect of ethnic origin could not be evaluated in any meaningful was because 98.8% of the population was Caucasian.
The trial excluded individuals with severe hepatic dysfunction; therefore, possible effects of severe hepatic decompensation on LY333334 disposition are not known. Within the range of liver-related parameters included in this analysis, there was no significant association between LY333334 clearance and values for serum bilirubin, alanine transaminase, aspartate transaminase or gamma glutamyl transferase.

There was no significant association between smoking status or alcohol use and LY333334 clearance or V/F.

Exposure estimates from the final population pharmacokinetic model for male patients in this study are presented for both doses in the following table:

LY333334 Treatment Group[a]	AUC (pg•hour/mL)	C_{max} (pg/mL)
20-µg		
Mean (%CV)	218.7 (24.9%)	122.4 (21.5%)
Median	208.6	121.2
5th and 95th Percentiles	148.2 - 321.0	87.0 - 165.1
n[b]	128	128
40-µg		
Mean (%CV)	434.4 (32.4%)	243.0 (27.3%)
Median	413.3	236.4
5th and 95th Percentiles	285.1 - 659.7	156.3 - 346.0
n[b]	121	121

Abbreviations: AUC = area under concentration-time curve; C_{max} = peak concentration; CV = coefficient of variation.

[a] Assigned treatment group at the time of a patient's first pharmacokinetic sample.

[b] n = Total number of patients included in the LY333334 pharmacokinetic analysis.

The sponsor investigated episodes of drug-related serious adverse events (i.e., events that were judged by the investigator to be "drug-related"), hypercalcemia, and hypercalciuria in patients with C_{max} or AUC values that were > the 95th percentile range for both the 20 µg and 40 µg doses. In the 20 µg group, patients had peak LY333334 concentrations > 165 pg/mL or AUC values > 320 pg.hour/mL. None of these patients had a drug-related serious adverse event during the study and none had a dosage reduction or discontinuation due to hypercalcemia or hypercalciuria.

A similar analysis for the 40 µg group showed that patients with LY333334 AUC or C_{max} in the 95th percentile had peak concentrations > 340 pg/mL or AUC values above 660 pg.hour/mL. None of these patients had a drug-related serious adverse event and one had a dosage reduction due to elevated serum and urine calcium. None of the patients discontinued due to hypercalcemia or hypercalciuria.