# CENTER FOR DRUG EVALUATION AND RESEARCH

***APPLICATION NUMBER:***

# 208743Orig1s000

# STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Science
Office of Biostatistics

# Statistical Review and Evaluation

## CARCINOGENICITY STUDY

| | |
|---|---|
| IND/NDA Number: | NDA-208743 |
| Drug Name: | BA058 |
| Indication: | Treatment of post-menopausal osteoporosis |
| Studies: | 104 Week Rat Carcinogenicity Study |
| Applicant: | Sponsor:<br>Radius Health, Inc.<br>201 Broadway, 6th Floor<br>Cambridge, MA 02139 USA<br>Testing Facility:<br>(b) (4) |
| Documents Reviewed: | Electronic submission: Submitted on June 15 2016<br>Electronic data: Submitted on June 15 2016 |
| Review Priority: | Standard |
| Biometrics Division: | Division of Biometrics - VI |
| Statistical Reviewer: | Hepei Chen |
| Concurring Reviewer: | Karl Lin, Ph.D. |
| Medical Division: | Division of Bone, Reproductive, and Urologic Products |
| Reviewing Pharmacologist: | Gemma Kuijpers, Ph.D. |
| Date Submitted: | Aug 11, 2016 |
| Keywords: | Carcinogenicity, Dose response |

# Table of Contents

## 1. Background

In this submission the sponsor included a carcinogenicity study report in rats. This study was intended to to determine the carcinogenic potential of BA058 a bone anabolic agent when given by daily subcutaneous injection for a minimum of 104 consecutive weeks to Fischer 344 rats and to provide data to support the use of BA058 in humans. Results of this review have been discussed with the reviewing pharmacologist Dr. Kuijpers.

In this review the phrase "dose response relationship" refers to the linear component (trend) of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor incidence rate as dose increases.

## 2. Study Design and Analysis

Two separate experiments, one in male rats and one in female rats were conducted. As indicated in Table 1, in each of these two experiments there were three treated groups, one vehicle control group, and one positive group. Three hundred Fischer 344 Albino rats of each sex were assigned randomly to the treated, control, and positive groups in equal size of 60 rats per group, except for 59 rats in mid dose group (Group 3) of male rats, and 61 rats in mid dose group (Group 3) of female rats. The dose levels for treated groups were 10, 25, and 50 µg/kg/day, respectively. In this review these dose groups were referred to as the low (Group 2), mid (Group 3), and high (Group 4) dose groups, respectively. The rats in the vehicle control group (Group 1) and the positive control group (Group 5) were administered with the reference item (0.9% sodium chloride for injection) and the positive item [Parathyroid hormone (PTH; 1-34)], respectively, and handled for the same duration and in the same manner as the treated groups.

**Table 1: Experimental Design in Rat Study**

| Group No. | No. of Toxicity Animals[a] | | Test Material | Dosae Level (µg/kg/day)[b] |
|---|---|---|---|---|
| | Male | Female | | |
| 1 | 60 | 60 | Vehicle control | 0 |
| 2 | 60 | 60 | BA058 low | 10 |
| 3 | 59 | 61 | BA058 mid | 25 |
| 4 | 60 | 60 | BA058 mid | 50 |
| 5 | 60 | 60 | Positive control | 30 |

a In this study, there was a pseudohermaprodite and the concerned animal found in Group 3 males has been reassigned to Group 3 females which ended up with 61 animals instead of 60.
b Given that the dosing periods for Groups 1, 2, 3, and 4 were not the same for all groups, the doses have been adjusted accordingly to (0, 9.85, 23.96, 47.31) for the males and to (0, 10, 25, 46.73) for the females.

Due to the decreased survival incidence reaching 15 animals per group (25%) in male rats, early terminations were conducted to the mid dose (Group 3), high dose (Group 4), and positive (Group 5) groups at Weeks 97, 88, and 99, respectively.

Throughout the study, animals were observed for general health/mortality and moribundity twice daily, once in the morning and once in the afternoon. The animals were removed from the cage, and a detailed clinical observation was performed weekly, beginning during Week -1. The presence of palpable masses was observed during the detailed examination. The site (location), size and appearance of any masses were recorded when first detected and, following this initial description, the presence or disappearance of any masses was monitored.

## 2.1. Sponsor's analyses

Given that the dosing periods for Groups 1, 2, 3, and 4 were not all the same, their corresponding arithmetic dose level scores (0, 10, 25, 50) were adjusted accordingly. The adjusted dose level scores given by (0, 9.85, 23.96, 47.31) for the males and by (0, 10, 25, 46.73) for the females were multiplied by 100 and the integer resulting scores were used to perform the overall trend test in the sponsor's analysis.

### 2.1.1. Survival analysis

The sponsor performed tests for dose response relationship and for each treated group (Group 2, 3, and 4) against the control group (Group 1) using Kaplan-Meier product-limit estimation curves, along with log-rank and Wilcoxon tests for each sex separately. The log-rank test was applied to the four groups in order to assess the significance of the overall group effect on mortality data, and was implemented in Proc Lifetest as a two-sided homogeneity test. The time to death or sacrifice was the dependent variable. The treatment group was included as the stratum. The animal was classified as uncensored when death occurred within the experimental period and was not accidental, while the animal with a death or sacrifice status recorded as a planned sacrifice (interim or terminal) or an accidental death was censored in the analysis.

If the log-rank test was found to be significant ($p \leq 0.05$), then the significance of a dose-related trend in mortality across all groups was evaluated using Tarone's method. Using the Multtest procedure (SAS/STAT), Tarone's test was implemented as a Peto two-sided test, with all uncensored deaths coded as 2 and all censored deaths coded as 0. Furthermore, the vehicle control group was compared to each of the three treated groups using a Peto two-sided test. When performing a pairwise comparison, only the data corresponding to the two compared groups were submitted to statistical analysis.

The sign of the statistic was used to indicate if either an increase or a decrease of the mortality rate across dose levels was observed for the trend test and each pairwise comparison. Each trend test and pairwise group comparison of interest was conducted at the 5% significance level. Significant trend and pairwise comparison results were reported as either $p \leq 0.001$, $p \leq 0.01$ or $p \leq 0.05$, where $p$ represented the obtained probability.

**Sponsor's findings**:

The sponsor's analysis showed that the numbers of rats surviving to their terminal necropsy were 20 (33%), 19 (32%), 15 (25%), 15 (25%), and 15 (25%) in Groups 1, 2, 3, 4, and 5 for male rats, respectively, and 27 (45%), 33 (55%), 25 (42%),16 (27%), and 26 (43%) in Groups 1, 2, 3, 4, and 5 for female rats, respectively.

For both male and female rats in the sponsor's analysis, the homogeneity comparison revealed significance among the three treated and the vehicle control groups. More precisely, this test revealed significance at the 5% level, with $p < 0.0001$ for the males and $p = 0.0007$ for the females. Consequently, for both males' and female' datasets, Tarone's trend test implemented as a Peto's two-sided test, was performed. The results of this test revealed significance, at the 5% level, of a dose-related trend in mortality, with $p < 0.0001$ for the males and $p = 0.0004$ for the females. The positive sign of the corresponding test statistics indicates that there is an increase of

the mortality rate across the three treated groups and the vehicle control group in both male and female rats.

Furthermore, the vehicle control group (Group 1) was compared to each of the other four groups (Group 2, 3, 4, and 5) using Peto's two-sided test. For the males, the results showed significance, at the 5% level, for the pairwise comparisons between Group 1 and each of the Groups 3, 4, and 5, with $p = 0.0021$, $p < 0.0001$, and $p = 0.0150$ respectively. However, for the females, the results showed significance, at the 5% level, only for the pairwise comparison between Groups 1 and 4, with $p = 0.0054$. For these four significant results, the positive sign of the corresponding test statistics indicates that the mortality rate in each of the respective group is higher than the one in the vehicle control group.

### 2.1.2. Tumor data analysis

For each sex, the statistical evaluation of tumor data were performed using the Multtest procedure of the SAS/STAT module and was limited, unle**ss** specified otherwise as described below, to all primary neoplastic lesions found in study plan-required tissues/sites, excluding metastatic tumors. The death time of all animals that died after the end of the experimental period were considered to be the first day of the scheduled terminal sacrifice period.

Neoplastic lesions listed under mammary gland and “Skin Combination”, were statistically analyzed, using all study animals, according to Peto’s onset rate method for tumors observed in a “mortality independent” context. The onset time of each palpable tumor was given by the first date of detection during the experimental period. The death time was used as the onset time for a lesion categorized as palpable and not detected in vivo. For combinations of palpable tumors, the earliest onset time was used in the analysis. For combinations of both palpable and non-palpable tumors, the animal death time was used for the analysis which was conducted using the “mortality dependent” method described hereafter for non-palpable tumors.

For the males, given that the sacrifice periods were not the same for Groups 1, 2, 3, and 4, three dataset were considered for the statistical analysis:

1. the first dataset consisted of datasets including Groups 1, 2, 3, and 4 with a common experimental period of 615 days (88 weeks),
2. the second dataset consisted of datasets including Groups 1, 2, and 3 with a common experimental period of 674 days (97 weeks),
3. the third dataset consisting of datasets including Groups 1 and 2 with a common experimental period of 721 days (103 weeks).

For the females, given that the sacrifice period is the same for Groups 1, 2, 3, and 4, datasets including the four groups with a common experimental period of 734 days (105 weeks) were used for the analysis.

A separate statistical analysis was conducted for each dataset containing the findings of each tumor type listed under a study plan-required tissue/site. However, for each tumor type tabulated under study plan-required “Skin” and under non study plan listed “skin miscellaneous” and “subcutaneous tissue”, the findings were combined together, and each of the resulting combinations were listed under a site denoted by “Skin Combination”. In addition, for each tumor type tabulated under study plan-required “Bones” and under non-study plan listed “Bone miscellaneous”, the findings were combined together, and each of the resulting combinations

was listed under a new site denoted by “Bone Combination”. The statistical analysis of each resulting combination dataset was based on all study animals. Furthermore, hemangiosarcomas were combined across each study plan-required tissue/site and each non-study plan-listed tissue/site under which they appear. The resulting combination was listed under a site denoted by “Hemangiosarcoma, all sites”. The statistical evaluation of the hemangiosarcomas was performed separately for each study plan-required tissue/site as well as for the site “Hemangiosarcoma, all sites”.

Using the derived outcomes from the processing of both fatal and incidental (non-palpable) tumors, a test statistic was built to perform a global survival-adjusted trend test on tumor data observed in a “mortality dependent” context. In this context, each neoplastic finding was classified as the cause of death (designated by the study pathologist as definitely or probably fatal) or not the cause of death (designated by the study pathologist as definitely or probably incidental or undetermined). All non-palpable neoplasms found in an animal were automatically classified as incidental if the animal in question died after the experimental period. For combinations of neoplastic findings, if an animal had one of the related tumors classified as fatal, then the neoplastic finding defined by this combination was classified as fatal. Otherwise, it was classified as incidental. Neoplastic findings classified as fatal and incidental were processed using the death rate method and the prevalence method respectively.

The processing of incidental tumors was done by creating a single separate interval for the time period following the experimental period (terminal sacrifice period) and by dividing the experimental period into the following fixed intervals:

1. For male rats first dataset, Days 1-364 (Weeks 1-51), Days 365-546 (Weeks 52-77), and Days 547-615 (Weeks 78-87),
2. For male rats second dataset Days 1-364 (Weeks 1-51), Days 365-546 (Weeks 52-77), Days 547-644 (Weeks 78-91), and Days 645-674 (Weeks 92-96),
3. For male rats third dataset, Days 1-364 (Weeks 1-51), Days 365-546 (Weeks 52-77), Days 547-644 (Weeks 78-91), and Days 645-721 (Weeks 92-102),
4. For female rats dataset, Days 1-364 (Weeks 1-51), Days 365-546 (Weeks 52-77), Days 547-644 (Weeks 78-91), and Days 645-734 (Weeks 92-104).

For the females’ dataset and for each of the first and second males’ datasets, the significance of a linear dose-related increase in tumor occurrence rates, across the vehicle control group (Group 1) and the three treated groups, was evaluated using Peto's survival-adjusted one-tailed trend test. In addition, pairwise groups comparisons were made using Peto’s one-sided trend test to determine if the tumor rate in each of the treated and positive groups is significantly higher than the tumor rate in the vehicle control group. The discrete permutation distribution was used to compute the corresponding $p$-value for each statistical test performed on a dataset containing 10 or less tumor occurrences.

**Adjustment for multiple testing:**

According to the FDA’s recommendations (2001), in the sponsor’s report, the dose-related increase in tumor incidence was considered significant if the $p$-value is less than or equal to 0.025 for rare tumors and less than or equal to 0.005 for common tumors (historical incidence of more than 1%); the increased tumor rate in a treated group when compared with the control group is considered significant when the $p$-value is less than or equal to 0.05 for a rare tumor, or is less than or equal to 0.01 for a common tumor.

**Sponsor's findings**:

A. Four groups (Group 1, 2, 3, and 4) for male and female rats

For the dataset including the four considered groups (Group 1, 2, 3, and 4), the sponsor analyzed the tumor data using a common experimental period of 615 days (87 weeks) for the male rats and 734 days (104 weeks) for the female rats. The corresponding adjusted arithmetic dose scores times 100, (0, 985, 2396, 4731) for the males and (0, 1000, 2500, 4673) for the females, were used to perform this overall trend test.

Based on FDA's recommendation, as shown in Table 2 below, the results of trend test corresponding to both osteoblastoma and osteosarcoma listed under bone combination for both male and female rats were considered to be statistically significant (all p-values < 0.005) regardless of the tumors type classification (rare or common). The results of the overall trend test corresponding to adenoma cortical listed under males' adrenal (p-value = 0.0157) and to fibroadenoma listed under females' injection site dorsal thoracic (p-value = 0.0053), were considered to be significant only if these tumors are classified as rare. The two remaining results of hemangioscarcoma under all sites hemangioscarcoma (p-value = 0.0337) and papilloma squamous cell under skin combination (p-value = 0.0323) in female rats were considered to be not statistically significant.

**Table 2: Analysis Results of Trend Test for Datasets Including Group 1, 2, 3, and 4.**

| Organ Name | Tumor Name | *P*-value | Flag |
|---|---|---|---|
| **---** ***MALES*** **---** | | | |
| Adrenal | Adenoma: cortical | 0.0157 | @ |
| Bone combination | Osteoblastoma | < 0.0001 | @@ |
| Bone combination | Osteosarcoma | < 0.0001 | @@ |
| **---** ***FEMALES*** **---** | | | |
| Bone combination | Osteoblastoma | 0.0034 | @@ |
| Bone combination | Osteosarcoma | < 0.0001 | @@ |
| Hemangiosarcoma, all sites | Hemangiosarcoma | 0.0337 | |
| Injection site dorsal thoracic, left | Fibroadenoma | 0.0053 | @ |
| Skin combination | Papilloma: squamous cell | 0.0323 | |

Based on FDA's recommendation, as shown in Table 3, all results of the pairwise comparisons between the treated groups and the vehicle control group corresponding to either osteoblastoma or osteosarcoma listed under bone combination for both male and female rats, were considered to be statistically significant (all p-values < 0.01) regardless of the tumors type classification (rare or common).

**Table 3: Analysis Results of Pairwise Comparisons for Datasets Including Group 1, 2, 3, and 4.**

| Organ Name | Tumor Name | Comparison | *P*-value | Flag |
|---|---|---|---|---|
| --- *MALES* --- | | | | |
| Bone combination | Osteoblastoma | 1 vs 3 | < 0.0001 | ## |
| Bone combination | Osteoblastoma | 1 vs 4 | < 0.0001 | ## |
| Bone combination | Osteosarcoma | 1 vs 2 | < 0.0001 | ## |
| Bone combination | Osteosarcoma | 1 vs 3 | < 0.0001 | ## |
| Bone combination | Osteosarcoma | 1 vs 4 | < 0.0001 | ## |
| --- *FEMALES* --- | | | | |
| Bone combination | Osteoblastoma | 1 vs 2 | 0.0050 | ## |
| Bone combination | Osteoblastoma | 1 vs 3 | 0.0058 | ## |
| Bone combination | Osteoblastoma | 1 vs 4 | 0.0014 | ## |
| Bone combination | Osteosarcoma | 1 vs 2 | 0.0022 | ## |
| Bone combination | Osteosarcoma | 1 vs 3 | < 0.0001 | ## |
| Bone combination | Osteosarcoma | 1 vs 4 | < 0.0001 | ## |

B. Three groups (Group 1, 2, and 3) for male rats

For the male dataset including the three considered groups (Group 1, 2, and 3), the sponsor analyzed the tumor data using a common experimental period of 674 days (Week 97). The corresponding adjusted arithmetic dose scores times 100, (0, 985, 2396), were used to perform this overall trend test.

Based on FDA's recommendation, as shown in Table 4, for both osteoblastoma and osteosarcoma listed under bone combination, the overall trend test were considered to be statistically significant (all p-values < 0.0001) regardless of the tumors type classification (rare or common). As shown in Table 5, the pairwise comparisons between the mid dose group (Group 3) and the vehicle control group (Group 1) for both osteoblastoma and osteosarcoma listed under bone combination, and between the low dose group (Group 2) and the vehicle control group (Group 1) for osteoblastoma listed under bone combination were considered to be statistically significant (all p-values < 0.0001) regardless of the tumors type classification (rare or common).

**Table 4: Analysis Results of Trend Test for Datasets Including Group 1, 2, and 3.**

| Organ Name | Tumor Name | *P*-value | Flag |
|---|---|---|---|
| ---- *MALES* --- | | | |
| Bone combination | Osteoblastoma | < 0.0001 | @@ |
| Bone combination | Osteosarcoma | < 0.0001 | @@ |

**Table 5: Analysis Results of Pairwise Comparisons for Datasets Including Group 1, 2, and 3.**

| Organ Name | Tumor Name | Comparison | *P*-value | Flag |
|---|---|---|---|---|
| --- *MALES* --- | | | | |
| Bone combination | Osteoblastoma | 1 vs 3 | <0.0001 | ## |
| Bone combination | Osteosarcoma | 1 vs 2 | <0.0001 | ## |
| Bone combination | Osteosarcoma | 1 vs 3 | <0.0001 | ## |

C. Three groups (Group 1 and 2) for male rats

For the male dataset including the low dose group (Group 2) and the vehicle control group (Groups 1) only, the sponsor performed the comparison between them using a common experimental period of 721 days (103 weeks). As shown in Table 6, a statistically significant

increase for osteossarcoma listed under bone combination (p-value < 0.0001) was noted regardless of the tumors type classification (rare or common).

**Table 6: Analysis Results of Comparisons for Datasets Including Group 1 and 2.**

| Organ Name | Tumor Name | Comparison | *P*-value | Flag |
|---|---|---|---|---|
| --- *MALES* --- | | | | |
| Bone combination | Osteosarcoma | 1 vs 2 | < 0.0001 | ## |

D. Positive control vs. vehicle control (Group 1 and 5) for male and female rats

The sponsor also performed the comparison between the vehicle control group (Group 1) and the positive control group (Group 5) using a common experimental period of 691 days (99 weeks) for the males and 734 (104 weeks) days for the females, respectively. As shown in Table 7, all results of the comparison between the vehicle control group (Group 1) and the positive control group (Group 5) were considered to be statistically significant (all p-values < 0.001) regardless of the tumors type classification.

**Table 7: Analysis Results of Comparisons for Datasets Including Group 1 and 5.**

| Organ Name | Tumor Name | Comparison | *P*-value | Flag |
|---|---|---|---|---|
| --- *MALES* --- | | | | |
| Bone combination | Osteoblastoma | 1 vs 5 | 0.0002 | ## |
| Bone combination | Osteosarcoma | 1 vs 5 | < 0.0001 | ## |
| --- FEMALES --- | | | | |
| Bone combination | Osteosarcoma | 1 vs 5 | < 0.0001 | ## |

## 2.2. Reviewer's analyses

To verify the sponsor's analyses and to perform additional analyses suggested by the reviewing toxicologist, this reviewer independently performed the survival and tumor data analyses using the data provided by the sponsor electronically.

According to the sponsor's report, the dosing periods for Groups 1, 2, 3, and 4 were not the same for all groups, their corresponding arithmetic dose level scores (0, 10, 25, 50) hence should be adjusted accordingly. However, because the detailed time information for each dose group cannot be located in the sponsor's report, the reviewing pharmacologist suggested using the adjusted arithmetic dose level scores used by the sponsor to perform the dose-response relationship test in the analysis. Therefore, the adjusted dose level scores of (0, 9.85, 23.96, and 47.31 for the males, 0, 10, 25, and 46.73 for the females) multiplied by 100 were used in this reviewer's report.

### 2.2.1. Survival analysis

The survival distributions of rats in all five groups (Groups 1, 2, 3, 4, and 5) were estimated using the Kaplan-Meier product limit method. The dose response relationship was tested across groups using the likelihood ratio test, and the homogeneity of survival distributions was tested using the log-rank test. The Kaplan-Meier curves for survival rates are given in Figures 1A and 1B in the appendix for all five groups in male and female rats, respectively. The intercurrent mortality data of all five groups, and the results of the tests for dose response relationship and homogeneity of

survivals for Groups 1, 2, 3, and 4 are given in Tables 1A and 1B in the appendix for male and female rats, respectively.

**Reviewer's findings**:

This reviewer's analysis showed that the numbers of rats surviving to their terminal necropsy were 18 (30.00%), 19 (31.67%), 15 (25.42%), 16 (26.67%), and 16 (26.67%) in Groups 1, 2, 3, 4, and 5 for male rats, respectively, and 26 (43.33%), 33 (55.00%), 24 (39.34%), 14 (23.33%), and 25 (41.67%) in Groups 1, 2, 3, 4, and 5 for female rats, respectively.

The reviewer's analysis showed a statistically significant positive dose-response relationship in mortality (p-value < 0.0001) along with statistically significant increases in the mid and high dose groups (Group 3 and 4) compared to the vehicle control group (Group 1) for the male rats (p-value = 0.0364 and <0.0001, respectively). Similarly, a statistically significant positive dose-response relationship in mortality (p-value = 0.0018) along with a statistically significant increase in the high dose group compared to the vehicle control group (p-value = 0.0028) were noted for the female rats.

Reviewer's comment: Based on the raw data (Tumor.xpt) provided by the sponsor, the numbers of rats surviving to their terminal necropsy for many dose groups reported in the sponsor's report (page 36) were incorrect. For example, for the high dose group, there were 16 and 14 animals with DTHSACST=2 (Terminal sacrifice) for the male and female rats, respectively, according to the raw data. However, the sponsor reported 15 and 16 (Group 4: 60-45=15 and 60-44=16) survivors in their report.

### 2.2.2. Tumor data analysis

The tumor data were analyzed for dose response relationships across Groups 1, 2, 3, and 4 and pairwise comparisons of each of the three treated groups (Groups 2, 3, and 4) and the positive control group (Group 5) against the vehicle control group (Group 1), using the Poly-k method described in the paper of Bailer and Portier (1988) and Bieler and Williams (1993).

In the ploy-k method, the adjustment for differences in mortality among treatment groups is made by modifying the number of animals at risk in the denominators in the calculations of overall tumor rates in the Cochran-Armitage test to reflect less-than-whole-animal contributions for animals that die without tumor before the end of the study (Bailer and Portier 1988). The modification is made by defining a new number of animals at risk for each treatment group. The number of animals at risk for the *i*-th treatment group $R^*_i$ is defined as $R^*_i = \sum w_{ij}$ where $w_{ij}$ is the weight for the *j*-th animal in the *i*-th treatment group, and the sum is over all animals in the group.

Bailer and Portier (1988) proposed the weight $w_{ij}$ as follows:

$w_{ij} = 1$ to animals dying with the tumor, and

$w_{ij} = ( t_{ij} / t_{sacr} )^k$ to animals dying without the tumor,

where $t_{ij}$ is the time of death of the *j*-th animal in the *i*-th treatment group, and $t_{sacr}$ is the planned (or intended) time of terminal sacrifice. The above formulas imply that animals living up to the end of the planned terminal sacrifice date without developing any tumor will also be assigned $w_{ij} = 1$ since $t_{ij} = t_{sacr}$.

Certain treatment groups of a study or the entire study may be terminated earlier than the planned (or intended) time of terminal sacrifice due to excessive mortalities. However, based on the principle of the Intention-to-treat (ITT) analysis in randomized trials, the $t_{sacr}$ should not be affected by the unplanned early terminations. The $t_{sacr}$ should always be equal to the planned (or intended) time of terminal sacrifice. For those animals that were sacrificed later than $t_{sacr}$, regardless their actual terminal sacrifice time, $t_{sacr}$ was used as their time of terminal sacrifice in the analysis.

One critical point for Poly-k test is the choice of the appropriate value of k, which depends on the tumor incidence pattern with the increased dose. For long term 104 week standard rat and mouse studies, a value of k=3 is suggested in the literature. Hence, this reviewer used k=3 for the analysis of this data.

The tumor rates and the p-values of the tested tumor types are listed in Tables 2A and 2B in the appendix for male and female rats, respectively.

**Adjustment for multiple testing:**

For the adjustment of multiple testing, this reviewer used the methodologies suggested in the FDA guidance for statistical design and analysis of carcinogenicity studies (2001). In order to keep the overall false-positive rate at the nominal level of approximately 10%, for both of the dose response relationship tests and the multiple pairwise comparisons of treated group with control group, the guidance suggests the use of a significance level $\alpha$=0.01 for common tumors and $\alpha$=0.05 for rare tumors (background rate of 1% or less) for a submission with one species,

**Reviewer's findings**:

The tumor types with p-values less than or equal to 0.05 for dose response relationship and/or pairwise comparisons of treated groups and combined control are reported in Table 8.

Based on the criteria of adjustment for multiple testing discussed previously, the results of the reviewer's analysis given in Table 8 below showed that for the combined osteoblastoma and osteosarcoma in various bone sites (femur, tibia, sternum, and vertebrae) and the whole body for both male and female rats, the dose-response relationships and the pairwise group comparisons between the treated groups and the vehicle control groups were statistically significant regardless of the tumors type classification (rare or common) for both male and female rats. The dose-response relationships of the adenoma cortical of the adrenal in male rats (p-value = 0.0287) and the fibroadenoma of the injection site dorsal thoracic in female rats (p-value = 0.0134), were considered to be significant only if these tumors are classified as rare. No other observed tumor types were noted to be statistically significant for the dose response relationships or pairwise comparisons in both male and female rats.

**Table 8. Summary Table of Tumor Types with P-Values ≤ 0.05 for Dose Response Relationship and/or Pairwise Comparisons of Treated Groups and Combined Control Group in Rats**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 30 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| **Male** | | | | | | |
| Adrenal | Adenoma: Cortical | 0/45 (59) | 0/43 (60) | 0/32 (59) | 2/25 (60) | 0/35 (60) |
| | | 0.0287 # | NC | NC | 0.1242 | NC |
| Bone-Femur | Osteoblastoma | 0/46 (60) | 0/43 (60) | 4/34 (59) | 4/27 (60) | 1/35 (60) |
| | | 0.0010 $ | NC | 0.0293 # | 0.0161 # | 0.4321 |
| | Osteosarcoma | 0/46 (60) | 4/44 (60) | 11/37 (59) | 18/34 (60) | 8/38 (60) |
| | | 0.0000 $ | 0.0531 | 0.0001 $ | 0.0000 $ | 0.0011 $ |
| | Osteoblastoma/Osteosarcoma | 0/46 (60) | 4/44 (60) | 15/39 (59) | 22/36 (60) | 9/38 (60) |
| | | 0.0000 $ | 0.0531 | 0.0000 $ | 0.0000 $ | 0.0004 $ |
| Bone-Sternum | Osteosarcoma | 0/46 (60) | 1/43 (60) | 5/35 (59) | 8/29 (60) | 5/38 (60) |
| | | 0.0000 $ | 0.4831 | 0.0127 # | 0.0003 $ | 0.0163 # |
| Bone-Tibia | Osteoblastoma | 0/46 (60) | 1/43 (60) | 10/37 (59) | 8/30 (60) | 4/36 (60) |
| | | 0.0000 $ | 0.4831 | 0.0001 $ | 0.0003 $ | 0.0337 $ |
| | Osteosarcoma | 1/46 (60) | 12/46 (60) | 21/41 (59) | 32/42 (60) | 14/41 (60) |
| | | 0.0000 $ | 0.0008 $ | 0.0000 $ | 0.0000 $ | 0.0001 $ |
| | Osteoblastoma/Osteosarcoma | 1/46 (60) | 13/46 (60) | 28/44 (59) | 37/45 (60) | 17/42 (60) |
| | | 0.0000# | 0.0004# | 0.0000# | 0.0000# | 0.0000# |
| Bone-Vertebrae L5- | Osteoblastoma | 0/46 (60) | 0/43 (60) | 4/34 (59) | 3/26 (60) | 3/36 (60) |
| | | 0.0037 $ | NC | 0.0293 # | 0.0436 # | 0.0806 |
| | Osteosarcoma | 0/46 (60) | 3/43 (60) | 3/34 (59) | 13/32 (60) | 7/39 (60) |
| | | 0.0000 $ | 0.1087 | 0.0728 | 0.0000 $ | 0.0031 $ |
| | Osteoblastoma/Osteosarcoma | 0/46 (60) | 3/43 (60) | 7/35 (59) | 16/33 (60) | 10/40 (60) |
| | | 0.0000 $ | 0.1087 | 0.0019 $ | 0.0000 $ | 0.0002 $ |
| Whole body | Osteoblastoma | 0/46 (60) | 1/43 (60) | 15/39 (59) | 20/37 (60) | 10/38 (60) |
| | | 0.0000 $ | 0.4831 | 0.0000 $ | 0.0000 $ | 0.0002 $ |
| | Osteosarcoma | 1/46 (60) | 31/52 (60) | 46/53 (59) | 52/54 (60) | 39/52 (60) |
| | | 0.0000 $ | 0.0000 $ | 0.0000 $ | 0.0000 $ | 0.0000 $ |
| | Osteosarcoma/Osteoblastoma | 1/46 (60) | 31/52 (60) | 48/54 (59) | 54/56 (60) | 42/53 (60) |
| | | 0.0000 $ | 0.0000 $ | 0.0000 $ | 0.0000 $ | 0.0000 $ |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;

$ = Statistically significant ($p<0.01$) for tests of dose response relationship or pairwise comparison regardless the type of tumor (rare or common);

# = Statistically significant ($p<0.05$) for tests of dose response relationship or pairwise comparison only if the tumor is considered to be rare;

NC = Not calculable.

**Table 8. Summary Table of Tumor Types with P-Values ≤ 0.05 for Dose Response Relationship and/or Pairwise Comparisons of Treated Groups and Combined Control Group in Rats (Continued)**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 0 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| **Female** | | | | | | |
| Bone-Femur | Osteoblastoma | 0/48 (60) | 2/50 (60) | 4/49 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.4467 | 0.2577 | 0.0612 | NC | NC |
| | Osteosarcoma | 1/48 (60) | 0/50 (60) | 3/49 (61) | 5/42 (60) | 4/48 (60) |
| | | 0.0057 $ | 1.0000 | 0.3164 | 0.0740 | 0.1808 |
| | Osteoblastoma/Osteosarcoma | 1/48 (60) | 2/50 (60) | 7/49 (61) | 5/42 (60) | 4/48 (60) |
| | | 0.0191 # | 0.5155 | 0.0317 # | 0.0740 | 0.1808 |
| Bone-Sternum | Osteoblastoma | 0/48 (60) | 0/49 (59) | 0/48 (61) | 0/41 (60) | 2/48 (60) |
| | | NC | NC | NC | NC | 0.2474 |
| | Osteosarcoma | 0/48 (60) | 1/49 (59) | 2/48 (61) | 6/44 (60) | 1/48 (60) |
| | | 0.0012 $ | 0.5052 | 0.2474 | 0.0099 $ | 0.5000 |
| | Osteoblastoma/Osteosarcoma | 0/48 (60) | 1/49 (59) | 2/48 (61) | 6/44 (60) | 3/48 (60) |
| | | 0.0012 $ | 0.5052 | 0.2474 | 0.0099 $ | 0.1211 |
| Bone-Tibia | Osteoblastoma | 0/48 (60) | 3/50 (60) | 0/48 (61) | 7/43 (60) | 1/48 (60) |
| | | 0.0016 $ | 0.1289 | NC | 0.0040 $ | 0.5000 |
| | Osteosarcoma | 0/48 (60) | 0/50 (60) | 1/48 (61) | 10/45 (60) | 6/49 (60) |
| | | 0.0000 $ | NC | 0.5000 | 0.0004 $ | 0.0142 $ |
| | Osteoblastoma/Osteosarcoma | 0/48 (60) | 3/50 (60) | 1/48 (61) | 16/46 (60) | 7/49 (60) |
| | | 0.0000 $ | 0.1289 | 0.5000 | 0.0000 $ | 0.0067 $ |
| Bone-Vertebrae L5- | Osteoblastoma | 0/48 (60) | 3/51 (60) | 3/48 (61) | 2/42 (60) | 1/48 (60) |
| | | 0.1940 | 0.1328 | 0.1211 | 0.2150 | 0.5000 |
| | Osteosarcoma | 0/48 (60) | 0/50 (60) | 3/49 (61) | 7/44 (60) | 4/48 (60) |
| | | 0.0001 $ | NC | 0.1250 | 0.0044 $ | 0.0586 |
| | Osteoblastoma/Osteosarcoma | 0/48 (60) | 3/51 (60) | 6/49 (61) | 9/45 (60) | 5/48 (60) |
| | | 0.0004 $ | 0.1328 | 0.0142 # | 0.0009 $ | 0.0280 # |
| I.S. Dors.Tho. Lt | Fibroadenoma | 0/48 (60) | 0/50 (60) | 2/48 (61) | 3/42 (60) | 1/48 (60) |
| | | 0.0134 # | NC | 0.2474 | 0.0977 | 0.5000 |
| Whole body | Osteoblastoma | 0/48 (60) | 8/51 (60) | 7/49 (61) | 9/44 (60) | 4/48 (60) |
| | | 0.0064 $ | 0.0037 $ | 0.0067 $ | 0.0008 $ | 0.0586 |
| | Osteosarcoma | 1/48 (60) | 11/52 (60) | 22/53 (61) | 37/55 (60) | 24/52 (60) |
| | | 0.0000 $ | 0.0030 $ | 0.0000 $ | 0.0000 $ | 0.0000 $ |
| | Osteosarcoma/Osteoblastoma | 1/48 (60) | 16/52 (60) | 27/53 (61) | 40/55 (60) | 27/52 (60) |
| | | 0.0000 $ | 0.0001 $ | 0.0000 $ | 0.0000 $ | 0.0000 $ |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;

$ = Statistically significant ($p<0.01$) for tests of dose response relationship or pairwise comparison regardless the type of tumor (rare or common);

# = Statistically significant ($p<0.05$) for tests of dose response relationship or pairwise comparison only if the tumor is considered to be rare;

NC = Not calculable.

## 3. Summary

In this submission the sponsor included a carcinogenicity study report in rats. This study was intended to to determine the carcinogenic potential of BA058 a bone anabolic agent when given by daily subcutaneous injection for a minimum of 104 consecutive weeks to Fischer 344 rats and to provide data to support the use of BA058 in humans.

Two separate experiments, one in male rats and one in female rats were conducted. In each of these two experiments there were three treated groups, one vehicle control group, and one positive group. Three hundred Fischer 344 Albino rats of each sex were assigned randomly to the treated, control, and positive groups in equal size of 60 rats per group, except for 59 rats in mid dose group (Group 3) of male rats, and 61 rats in mid dose group (Group 3) of female rats. The dose levels for treated groups were 10, 25, and 50 µg/kg/day, respectively.

Due to the decreased survival incidence reaching 15 animals per group (25%) in male rats, early terminations were conducted to the mid dose (Group 3), high dose (Group 4), and positive (Group 5) groups at Weeks 97, 88, and 99, respectively.

This reviewer's analysis showed that the numbers of rats surviving to their terminal necropsy were 18 (30.00%), 19 (31.67%), 15 (25.42%), 16 (26.67%), and 16 (26.67%) in Groups 1, 2, 3, 4, and 5 for male rats, respectively, and 26 (43.33%), 33 (55.00%), 24 (39.34%), 14 (23.33%), and 25 (41.67%) in Groups 1, 2, 3, 4, and 5 for female rats, respectively.

The reviewer's analysis showed a statistically significant positive dose-response relationship in mortality (p-value < 0.0001) along with statistically significant increases in the mid and high dose groups (Group 3 and 4) compared to the vehicle control group (Group 1) for the male rats (p-value = 0.0364 and <0.0001, respectively). Similarly, a statistically significant positive dose-response relationship in mortality (p-value = 0.0018) along with a statistically significant increase in the high dose group compared to the vehicle control group (p-value = 0.0028) were noted for the female rats.

Based on the criteria of adjustment for multiple testing discussed previously, the results of the reviewer's analysis given showed that for the combined osteoblastoma and osteosarcoma in various bone sites (femur, tibia, sternum, and vertebrae) and the whole body for both male and female rats, the dose-response relationships and the pairwise group comparisons between the treated groups and the vehicle control groups were statistically significant regardless of the tumors type classification (rare or common) for both male and female rats. The dose-response relationships of the adenoma cortical of the adrenal in male rats (p-value = 0.0287) and the fibroadenoma of the injection site dorsal thoracic in female rats (p-value = 0.0134), were considered to be significant only if these tumors are classified as rare. No other observed tumor types were noted to be statistically significant for the dose response relationships or pairwise comparisons in both male and female rats.

Hepei Chen.
Mathematical Statistician

Concur: Karl Lin, Ph.D.
Team Leader, DBVI

Cc: Archival NDA-208743

Dr. Gemma Kuijpers
Dr. Lillian Patrician
Dr. Mohammad Atiar Rahman

# 4. Appendix

## Table 1A: Intercurrent Mortality Rate in Male Rats

| | Positive Control | | Vehicle Control | | Low | | Mid | | High | |
|---|---|---|---|---|---|---|---|---|---|---|
| Week / Type of Death | No. of Death | No. of Death | Cum % | No. of Death | Cum % | No. of Death | Cum % | No. of Death | Cum % | Cum % |
| 0 - 52 | 3 | 5.00 | | | 3 | 5.00 | 6 | 10.17 | 4 | 6.67 |
| 53 - 78 | 16 | 31.67 | 8 | 16.67 | 4 | 11.67 | 12 | 30.51 | 25 | 48.33 |
| 79 - 91 | 16 | 58.33 | 13 | 38.33 | 19 | 43.33 | 20 | 64.41 | 15 | 73.33 |
| 92 - 104 | 9 | 73.33 | 19 | 70.00 | 15 | 68.33 | 6 | 74.58 | | |
| 105 | | | 2 | 3.33 | | | | | | |
| Terminal sacrifice | 16 | 26.67 | 18 | 30.00 | 19 | 31.67 | 15 | 25.42 | 16 | 26.67 |
| Total | 60 | | 60 | | 60 | | 59 | | 60 | |

| Test | All Dose Groups | Vehicle Control vs. Low | Vehicle Control vs. Mid | Vehicle Control vs. High |
|---|---|---|---|---|
| Dose-Response (Likelihood Ratio) | <.0001** | 0.4046 | 0.0447* | <.0001** |
| Homogeneity (Log-Rank) | <.0001** | 0.4008 | 0.0364* | <.0001** |

#All Cum. % Cumulative Percentage except for Terminal sacrifice;
* = Significant at 5% level; ** = Significant at 1% level.

## Table 1B: Intercurrent Mortality Rate in Female Rats

| | Positive Control | | Vehicle Control | | Low | | Mid | | High | |
|---|---|---|---|---|---|---|---|---|---|---|
| Week / Type of Death | No. of Death | No. of Death | Cum % | No. of Death | Cum % | No. of Death | Cum % | No. of Death | Cum % | Cum % |
| 0 - 52 | | | 1 | 3.33 | 2 | 3.33 | 2 | 4.92 | 1 | 5.00 |
| 53 - 78 | 9 | 16.67 | 7 | 15.00 | 2 | 6.67 | 6 | 14.75 | 11 | 23.33 |
| 79 - 91 | 11 | 35.00 | 6 | 25.00 | 9 | 21.67 | 11 | 32.79 | 17 | 51.67 |
| 92 - 104 | 14 | 58.33 | 19 | 56.67 | 13 | 43.33 | 17 | 60.66 | 15 | 76.67 |
| 92 - 104 | 1 | 1.67 | 1 | 1.67 | | | 1 | 1.64 | 2 | 3.33 |
| Accidental Death | | | | | 1 | 1.67 | | | | |
| Terminal sacrifice | 25 | 41.67 | 26 | 43.33 | 33 | 55.00 | 24 | 39.34 | 14 | 23.33 |
| Total | 60 | | 60 | | 60 | | 61 | | 60 | |

| Test | All Dose Groups | Vehicle Control vs. Low | Vehicle Control vs. Mid | Vehicle Control vs. High |
|---|---|---|---|---|
| Dose-Response (Likelihood Ratio) | 0.0018** | 0.0955 | 0.7631 | 0.0044** |
| Homogeneity (Log-Rank) | 0.0007** | 0.0984 | 0.7595 | 0.0028** |

#All Cum. % Cumulative Percentage except for Terminal sacrifice;
** = Significant at 1% level.

**Table 2A: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Male Rats**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 0 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| Adrenal | Adenoma: Cortical | 0/45 (59) | 0/43 (60) | 0/32 (59) | 2/25 (60) | 0/35 (60) |
| | | 0.0287 | NC | NC | 0.1242 | NC |
| | Benign Pheochromocytoma | 10/47 (59) | 8/45 (60) | 7/35 (59) | 3/26 (60) | 4/36 (60) |
| | | 0.8158 | 0.7529 | 0.6585 | 0.9170 | 0.9385 |
| | Benign Pheochromocytoma: Complex | 1/45 (59) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 0/35 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| | Benign Pheochromocytoma/ Benign Pheochromocytoma: Complex | 11/48 (59) | 8/45 (60) | 7/35 (59) | 3/26 (60) | 4/36 (60) |
| | | 0.8555 | 0.8079 | 0.7195 | 0.9383 | 0.9570 |
| | Malignant Pheochromocytoma | 1/45 (59) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 0/35 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| | Benign Pheochromocytoma/ Benign Pheochromocytoma: Complex/ Malignant Pheochromocytoma | 12/48 (59) | 8/45 (60) | 7/35 (59) | 3/26 (60) | 4/36 (60) |
| | | 0.8957 | 0.8644 | 0.7870 | 0.9581 | 0.9730 |
| Bone Miscellaneous | Chordoma: Malignant | 0/2 (4) | 0/25 (35) | 0/26 (46) | 1/21 (47) | 2/26 (41) |
| | | 0.2838 | NC | NC | 0.9130 | 0.1402 |
| | Hemangioma | 0/2 (4) | 1/26 (35) | 0/26 (46) | 0/21 (47) | 0/25 (41) |
| | | 0.6267 | 0.9286 | NC | NC | NC |
| | Hemangiosarcoma | 1/3 (4) | 0/25 (35) | 1/27 (46) | 0/21 (47) | 0/25 (41) |
| | | 0.8389 | 0.8929 | 0.8069 | 0.8750 | 0.1071 |
| | Hemangioma/Hemangiosarcoma | 1/3 (4) | 1/26 (35) | 1/27 (46) | 0/21 (47) | 0/25 (41) |
| | | 0.9202 | 0.8005 | 0.8069 | 0.8750 | 0.1071 |
| | Osteoblastoma | 0/2 (4) | 0/25 (35) | 0/26 (46) | 8/25 (47) | 3/25 (41) |
| | | 0.0000$ | NC | NC | 0.4872 | 0.2137 |
| | Osteoma | 0/2 (4) | 0/25 (35) | 0/26 (46) | 0/21 (47) | 1/25 (41) |
| | | NC | NC | NC | NC | 0.0741 |
| | Osteosarcoma | 0/2 (4) | 19/31 (35) | 31/41 (46) | 29/37 (47) | 30/38 (41) |
| | | 0.0318 | 0.1723 | 0.0731 | 0.0607 | 0.9423 |
| Bone-Femur | Osteoblastoma | 0/46 (60) | 0/43 (60) | 4/34 (59) | 4/27 (60) | 1/35 (60) |
| | | 0.0010 $ | NC | 0.0293 $ | 0.0161 $ | 0.4321 |
| | Osteosarcoma | 0/46 (60) | 4/44 (60) | 11/37 (59) | 18/34 (60) | 8/38 (60) |
| | | 0.0000 $ | 0.0531 | 0.0001 $ | 0.0000 $ | 0.0011 $ |
| | Osteoblastoma/Osteosarcoma | 0/46 (60) | 4/44 (60) | 15/39 (59) | 22/36 (60) | 9/38 (60) |
| | | 0.0000 $ | 0.0531 | 0.0000 $ | 0.0000 $ | 0.0004 $ |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

**Table 2A: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Male Rats (Continued)**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 0 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| Bone-Sternum | Osteosarcoma | 0/46 (60) | 1/43 (60) | 5/35 (59) | 8/29 (60) | 5/38 (60) |
| | | 0.0000 $ | 0.4831 | 0.0127 $ | 0.0003 $ | 0.0163 $ |
| Bone-Tibia | Osteoblastoma | 0/46 (60) | 1/43 (60) | 10/37 (59) | 8/30 (60) | 4/36 (60) |
| | | 0.0000 $ | 0.4831 | 0.0001 $ | 0.0003 $ | 0.0337 $ |
| | Osteosarcoma | 1/46 (60) | 12/46 (60) | 21/41 (59) | 32/42 (60) | 14/41 (60) |
| | | 0.0000# | 0.0008# | 0.0000# | 0.0000# | 0.0001# |
| | Osteoblastoma/Osteosarcoma | 1/46 (60) | 13/46 (60) | 28/44 (59) | 37/45 (60) | 17/42 (60) |
| | | 0.0000# | 0.0004# | 0.0000# | 0.0000# | 0.0000# |
| Bone-Vertebrae L5- | Osteoblastoma | 0/46 (60) | 0/43 (60) | 4/34 (59) | 3/26 (60) | 3/36 (60) |
| | | 0.0037 $ | NC | 0.0293 $ | 0.0436 $ | 0.0806 |
| | Osteosarcoma | 0/46 (60) | 3/43 (60) | 3/34 (59) | 13/32 (60) | 7/39 (60) |
| | | 0.0000 $ | 0.1087 | 0.0728 | 0.0000 $ | 0.0031 $ |
| | Osteoblastoma/Osteosarcoma | 0/46 (60) | 3/43 (60) | 7/35 (59) | 16/33 (60) | 10/40 (60) |
| | | 0.0000 $ | 0.1087 | 0.0019 $ | 0.0000 $ | 0.0002 $ |
| Brain | Malignant Granular Cell Tumor | 1/45 (59) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 0/35 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| | Malignant Mixed Glioma | 0/45 (59) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 1/36 (60) |
| | | NC | NC | NC | NC | 0.4444 |
| Epididymis | Mesothelioma(M) | 4/47 (60) | 0/43 (60) | 3/34 (59) | 1/25 (60) | 1/35 (60) |
| | | 0.5745 | 1.0000 | 0.6294 | 0.8904 | 0.9438 |
| Esophagus | Leiomyoma | 0/46 (60) | 0/43 (60) | 1/33 (59) | 0/25 (60) | 0/35 (60) |
| | | 0.3946 | NC | 0.4177 | NC | NC |
| Harderian Gland | Adenoma | 0/46 (60) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 1/36 (60) |
| | | NC | NC | NC | NC | 0.4390 |
| Hemolymph Tissue | Leukemia: Large Granular Lymphocyte | 36/56 (60) | 15/47 (60) | 3/34 (59) | 2/26 (60) | 9/39 (60) |
| | | 1.0000 | 0.9998 | 1.0000 | 1.0000 | 1.0000 |
| | Malignant Lymphoma | 0/46 (60) | 0/43 (60) | 1/33 (59) | 0/25 (60) | 0/35 (60) |
| | | 0.3946 | NC | 0.4177 | NC | NC |
| I.S. Dors.Tho. Lt | Sarcoma (Not Otherwise Specified) | 0/46 (60) | 1/43 (60) | 0/32 (59) | 0/25 (60) | 0/35 (60) |
| | | 0.6849 | 0.4831 | NC | NC | NC |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

**Table 2A: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Male Rats (Continued)**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 0 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| I.S. Dors.Tho. Rt | Adenoma: Basal Cell | 1/46 (60) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 0/35 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| | Fibroma | 0/46 (60) | 2/43 (60) | 2/33 (59) | 0/25 (60) | 1/35 (60) |
| | | 0.4478 | 0.2306 | 0.1714 | NC | 0.4321 |
| | Fibrosarcoma | 0/46 (60) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 1/35 (60) |
| | | NC | NC | NC | NC | 0.4321 |
| | Fibroma/Fibrosarcoma | 0/46 (60) | 2/43 (60) | 2/33 (59) | 0/25 (60) | 2/35 (60) |
| | | 0.4478 | 0.2306 | 0.1714 | NC | 0.1836 |
| | Keratoacanthoma | 2/46 (60) | 0/43 (60) | 0/32 (59) | 1/25 (60) | 0/35 (60) |
| | | 0.5298 | 1.0000 | 1.0000 | 0.7344 | 1.0000 |
| I.S. Interscapular | Lipoma | 0/46 (60) | 1/43 (60) | 0/32 (59) | 0/25 (60) | 0/35 (60) |
| | | 0.6849 | 0.4831 | NC | NC | NC |
| I.S. Lumbar, Left | Adenoma: Basal Cell | 0/46 (60) | 0/43 (60) | 1/33 (59) | 0/25 (60) | 0/35 (60) |
| | | 0.3946 | NC | 0.4177 | NC | NC |
| | Keratoacanthoma | 0/46 (60) | 0/43 (60) | 0/32 (59) | 1/25 (60) | 0/35 (60) |
| | | 0.1712 | NC | NC | 0.3521 | NC |
| I.S. Lumbar, Right | Fibroma | 1/46 (60) | 0/43 (60) | 1/33 (59) | 0/25 (60) | 1/35 (60) |
| | | 0.6351 | 1.0000 | 0.6641 | 1.0000 | 0.6806 |
| I.S. Sacral, Left | Adenoma: Basal Cell | 0/46 (60) | 0/43 (60) | 0/32 (59) | 1/25 (60) | 1/36 (60) |
| | | 0.1712 | NC | NC | 0.3521 | 0.4390 |
| | Fibroma | 0/46 (60) | 1/43 (60) | 0/32 (59) | 0/25 (60) | 1/36 (60) |
| | | 0.6849 | 0.4831 | NC | NC | 0.4390 |
| | Keratoacanthoma | 0/46 (60) | 0/43 (60) | 0/32 (59) | 1/25 (60) | 0/35 (60) |
| | | 0.1712 | NC | NC | 0.3521 | NC |
| Kidney | Adenoma: Tubular Cell | 0/46 (60) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 1/35 (60) |
| | | NC | NC | NC | NC | 0.4321 |
| Liver | Adenoma: Hepatocellular | 4/46 (60) | 1/44 (60) | 0/32 (59) | 0/25 (60) | 3/36 (60) |
| | | 0.9974 | 0.9688 | 1.0000 | 1.0000 | 0.6701 |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

**Table 2A: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Male Rats (Continued)**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 0 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| Lung | Adenoma: Alveolar/Bronchiolar | 3/46 (60) | 3/43 (60) | 0/32 (59) | 1/25 (60) | 0/35 (60) |
| | | 0.8204 | 0.6289 | 1.0000 | 0.8321 | 1.0000 |
| | Carcinoma:Alveolar/Bronchiolar | 1/46 (60) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 0/35 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| | Adenoma: Alveolar/Bronchiolar/ Carcinoma:Alveolar/Bronchiolar | 4/46 (60) | 3/43 (60) | 0/32 (59) | 1/25 (60) | 0/35 (60) |
| | | 0.8945 | 0.7541 | 1.0000 | 0.8947 | 1.0000 |
| | Carcinoma: Squamous Cell | 0/46 (60) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 1/36 (60) |
| | | NC | NC | NC | NC | 0.4390 |
| Mammary Gland | Fibroadenoma | 2/42 (54) | 0/38 (53) | 0/28 (52) | 0/22 (54) | 0/33 (55) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| Pancreas | Adenoma: Islet Cell | 3/46 (60) | 1/43 (60) | 2/33 (59) | 0/25 (60) | 3/36 (60) |
| | | 0.8400 | 0.9332 | 0.7002 | 1.0000 | 0.5387 |
| | Carcinoma: Islet Cell | 1/46 (60) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 1/35 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.6806 |
| | Adenoma: Islet Cell/ Carcinoma: Islet Cell | 4/46 (60) | 1/43 (60) | 2/33 (59) | 0/25 (60) | 4/36 (60) |
| | | 0.9124 | 0.9670 | 0.8036 | 1.0000 | 0.4984 |
| Parathyroid Gland | Adenoma | 1/41 (55) | 0/33 (47) | 0/26 (46) | 0/16 (38) | 0/26 (44) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| Pituitary | Adenoma: Pars Distalis | 30/51 (59) | 35/50 (59) | 23/40 (58) | 19/33 (58) | 28/44 (60) |
| | | 0.7015 | 0.1674 | 0.6339 | 0.6338 | 0.3946 |
| | Ganglioneuroma | 0/45 (59) | 0/43 (59) | 0/32 (58) | 0/23 (58) | 1/36 (60) |
| | | NC | NC | NC | NC | 0.4444 |
| Prostate | Adenoma | 0/46 (60) | 0/43 (60) | 0/32 (57) | 1/25 (60) | 0/35 (60) |
| | | 0.1712 | NC | NC | 0.3521 | NC |
| Seminal Vesicle | Adenoma | 0/46 (60) | 0/43 (60) | 1/32 (58) | 1/25 (60) | 0/35 (60) |
| | | 0.1039 | NC | 0.4103 | 0.3521 | NC |
| Skin | Adenoma: Basal Cell | 0/46 (60) | 0/43 (60) | 1/33 (59) | 0/25 (60) | 0/35 (60) |
| | | 0.3946 | NC | 0.4177 | NC | NC |
| Skin Miscellaneous | Carcinoma: Squamous Cell | 0/32 (43) | 1/32 (42) | 1/22 (36) | 0/16 (35) | 1/30 (46) |
| | | 0.4719 | 0.5000 | 0.4074 | NC | 0.4839 |
| | Papilloma: Squamous Cell | 0/32 (43) | 0/32 (42) | 0/22 (36) | 1/17 (35) | 0/29 (46) |
| | | 0.1650 | NC | NC | 0.3469 | NC |
| | Keratoacanthoma | 2/32 (43) | 0/32 (42) | 0/22 (36) | 1/17 (35) | 1/29 (46) |
| | | 0.5136 | 1.0000 | 1.0000 | 0.7308 | 0.8622 |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

**Table 2A: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Male Rats (Continued)**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 0 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| Spleen | Fibroma | 0/46 (60) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 1/35 (60) |
| | | NC | NC | NC | NC | 0.4321 |
| | Fibrosarcoma | 0/46 (60) | 1/43 (60) | 0/32 (59) | 0/25 (60) | 1/35 (60) |
| | | 0.6849 | 0.4831 | NC | NC | 0.4321 |
| | Fibroma/Fibrosarcoma | 0/46 (60) | 1/43 (60) | 0/32 (59) | 0/25 (60) | 2/35 (60) |
| | | 0.6849 | 0.4831 | NC | NC | 0.1836 |
| | Hemangiosarcoma | 0/46 (60) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 1/35 (60) |
| | | NC | NC | NC | NC | 0.4321 |
| Stomach | Papilloma: Squamous Cell | 0/46 (60) | 0/43 (60) | 0/32 (59) | 0/25 (60) | 1/35 (60) |
| | | NC | NC | NC | NC | 0.4321 |
| Testis | Adenoma: Interstitial Cell | 44/53 (60) | 38/51 (60) | 35/45 (59) | 37/44 (60) | 35/48 (60) |
| | | 0.3446 | 0.9038 | 0.8187 | 0.5550 | 0.9293 |
| Thyroid | Adenoma: C-Cell | 9/47 (60) | 8/44 (60) | 9/35 (59) | 5/27 (60) | 6/37 (60) |
| | | 0.4255 | 0.6497 | 0.3279 | 0.6397 | 0.7353 |
| | Adenoma: Follicular Cell | 3/46 (60) | 3/44 (60) | 1/33 (59) | 0/25 (60) | 1/35 (60) |
| | | 0.9421 | 0.6401 | 0.8914 | 1.0000 | 0.9019 |
| | Carcinoma: Follicular Cell | 1/46 (60) | 0/43 (60) | 1/32 (59) | 0/25 (60) | 0/35 (60) |
| | | 0.6300 | 1.0000 | 0.6553 | 1.0000 | 1.0000 |
| | Adenoma: Follicular Cell/ Carcinoma: Follicular Cell | 4/46 (60) | 3/44 (60) | 2/33 (59) | 0/25 (60) | 1/35 (60) |
| | | 0.9356 | 0.7641 | 0.8036 | 1.0000 | 0.9465 |
| Tongue | Benign Schwannoma | 0/46 (60) | 1/43 (60) | 0/32 (59) | 0/25 (60) | 0/35 (60) |
| | | 0.6849 | 0.4831 | NC | NC | NC |
| Whole body | Osteoblastoma | 0/46 (60) | 1/43 (60) | 15/39 (59) | 20/37 (60) | 10/38 (60) |
| | | 0.0000 $ | 0.4831 | 0.0000 $ | 0.0000 $ | 0.0002 $ |
| | Osteosarcoma | 1/46 (60) | 31/52 (60) | 46/53 (59) | 52/54 (60) | 39/52 (60) |
| | | 0.0000# | 0.0000# | 0.0000# | 0.0000# | 0.0000# |
| | Osteosarcoma/Osteoblastoma | 1/46 (60) | 31/52 (60) | 48/54 (59) | 54/56 (60) | 42/53 (60) |
| | | 0.0000# | 0.0000# | 0.0000# | 0.0000# | 0.0000# |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

**Table 2B: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Female Rats**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 0 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| Adrenal | Adenoma: Cortical | 1/48 (60) | 1/50 (60) | 1/49 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.8097 | 0.7627 | 0.7577 | 1.0000 | 1.0000 |
| | Benign Pheochromocytoma | 4/49 (60) | 2/50 (60) | 1/49 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.9890 | 0.9024 | 0.9719 | 1.0000 | 1.0000 |
| Bone-Femur | Osteoblastoma | 0/48 (60) | 2/50 (60) | 4/49 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.4467 | 0.2577 | 0.0612 | NC | NC |
| | Osteosarcoma | 1/48 (60) | 0/50 (60) | 3/49 (61) | 5/42 (60) | 4/48 (60) |
| | | 0.0057 | 1.0000 | 0.3164 | 0.0740 | 0.1808 |
| | Osteoblastoma/Osteosarcoma | 1/48 (60) | 2/50 (60) | 7/49 (61) | 5/42 (60) | 4/48 (60) |
| | | 0.0191 | 0.5155 | 0.0317 | 0.0740 | 0.1808 |
| Bone Miscellaneous | Chordoma: Malignant | 0/3 (4) | 0/17 (19) | 0/22 (26) | 1/26 (37) | 0/22 (26) |
| | | 0.3824 | NC | NC | 0.8966 | NC |
| | Fibrosarcoma | 0/3 (4) | 0/17 (19) | 1/22 (26) | 0/25 (37) | 1/22 (26) |
| | | 0.3731 | NC | 0.8800 | NC | 0.1200 |
| | Hemangiosarcoma | 0/3 (4) | 0/17 (19) | 0/22 (26) | 1/25 (37) | 0/22 (26) |
| | | 0.3731 | NC | NC | 0.8929 | NC |
| | Osteoblastoma | 0/3 (4) | 0/17 (19) | 0/22 (26) | 1/26 (37) | 0/22 (26) |
| | | 0.3824 | NC | NC | 0.8966 | NC |
| | Osteoma | 0/3 (4) | 0/17 (19) | 1/22 (26) | 0/25 (37) | 0/22 (26) |
| | | 0.3731 | NC | 0.8800 | NC | NC |
| | Osteosarcoma | 0/3 (4) | 10/19 (19) | 17/25 (26) | 23/33 (37) | 13/24 (26) |
| | | 0.0418 | 0.1429 | 0.0504 | 0.0401 $ | 0.8756 |
| | Osteoblastoma/Osteoma | 0/3 (4) | 0/17 (19) | 1/22 (26) | 1/26 (37) | 0/22 (26) |
| | | 0.3938 | NC | 0.8800 | 0.8966 | NC |
| Bone-Sternum | Osteoblastoma | 0/48 (60) | 0/49 (59) | 0/48 (61) | 0/41 (60) | 2/48 (60) |
| | | NC | NC | NC | NC | 0.2474 |
| | Osteosarcoma | 0/48 (60) | 1/49 (59) | 2/48 (61) | 6/44 (60) | 1/48 (60) |
| | | 0.0012 $ | 0.5052 | 0.2474 | 0.0099 $ | 0.5000 |
| | Osteoblastoma/Osteosarcoma | 0/48 (60) | 1/49 (59) | 2/48 (61) | 6/44 (60) | 3/48 (60) |
| | | 0.0012 $ | 0.5052 | 0.2474 | 0.0099 $ | 0.1211 |
| Bone-Tibia | Osteoblastoma | 0/48 (60) | 3/50 (60) | 0/48 (61) | 7/43 (60) | 1/48 (60) |
| | | 0.0016 $ | 0.1289 | NC | 0.0040 $ | 0.5000 |
| | Osteosarcoma | 0/48 (60) | 0/50 (60) | 1/48 (61) | 10/45 (60) | 6/49 (60) |
| | | 0.0000 $ | NC | 0.5000 | 0.0004 $ | 0.0142 $ |
| | Osteoblastoma/Osteosarcoma | 0/48 (60) | 3/50 (60) | 1/48 (61) | 16/46 (60) | 7/49 (60) |
| | | 0.0000 $ | 0.1289 | 0.5000 | 0.0000 $ | 0.0067 $ |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

**Table 2B: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Female Rats (Continued)**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 0 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| Bone-Vertebrae L5- | Osteoblastoma | 0/48 (60) | 3/51 (60) | 3/48 (61) | 2/42 (60) | 1/48 (60) |
| | | 0.1940 | 0.1328 | 0.1211 | 0.2150 | 0.5000 |
| | Osteosarcoma | 0/48 (60) | 0/50 (60) | 3/49 (61) | 7/44 (60) | 4/48 (60) |
| | | 0.0001 $ | NC | 0.1250 | 0.0044 $ | 0.0586 |
| | Osteoblastoma/Osteosarcoma | 0/48 (60) | 3/51 (60) | 6/49 (61) | 9/45 (60) | 5/48 (60) |
| | | 0.0004 $ | 0.1328 | 0.0142 $ | 0.0009 $ | 0.0280 $ |
| Brain | Malignant Astrocytoma | 0/48 (60) | 1/50 (60) | 1/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.5942 | 0.5102 | 0.5000 | NC | NC |
| | Malignant Oligodendroglioma | 0/48 (60) | 1/51 (60) | 0/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.7447 | 0.5152 | NC | NC | NC |
| Cervix | Benign Granular Cell Tumor | 2/48 (60) | 0/50 (60) | 0/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| | Fibroma | 0/48 (60) | 0/50 (60) | 0/48 (61) | 1/42 (60) | 0/48 (60) |
| | | 0.2234 | NC | NC | 0.4667 | NC |
| | Leiomyosarcoma | 1/48 (60) | 0/50 (60) | 1/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.7267 | 1.0000 | NC | 1.0000 | 1.0000 |
| | Polyp: Stromal | 1/48 (60) | 0/50 (60) | 1/49 (61) | 0/41 (60) | 1/48 (60) |
| | | 0.7296 | 1.0000 | 0.7577 | 1.0000 | NC |
| | Sarcoma: Endometrial Stromal | 0/48 (60) | 0/50 (60) | 0/48 (61) | 1/41 (60) | 0/48 (60) |
| | | 0.2193 | NC | NC | 0.4607 | NC |
| | Polyp: Stromal/ Sarcoma: Endometrial Stromal | 1/48 (60) | 0/50 (60) | 1/49 (61) | 1/41 (60) | 1/48 (60) |
| | | 0.3622 | 1.0000 | 0.7577 | 0.7120 | NC |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

**Table 2B: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Female Rats (Continued)**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 0 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| Heart | Benign Schwannoma: Intramural | 1/48 (60) | 0/50 (60) | 0/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| Hemolym. Tissue | Histiocytic Sarcoma | 1/48 (60) | 0/50 (60) | 0/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| | Leukemia: Large Granular Lymphocyte | 19/53 (60) | 12/53 (60) | 19/56 (61) | 16/45 (60) | 21/53 (60) |
| | | 0.3106 | 0.9566 | 0.6596 | 0.5948 | 0.4207 |
| | Malignant Lymphoma | 0/48 (60) | 1/50 (60) | 0/48 (61) | 1/42 (60) | 0/48 (60) |
| | | 0.2831 | 0.5102 | NC | 0.4667 | NC |
| | Leukemia: Large Granular Lymphocyte/ Malignant Lymphoma | 19/53 (60) | 13/53 (60) | 19/56 (61) | 17/46 (60) | 21/53 (60) |
| | | 0.2863 | 0.9310 | 0.6596 | 0.5373 | 0.4207 |
| I.S. Dors.Tho. Lt | Fibroadenoma | 0/48 (60) | 0/50 (60) | 2/48 (61) | 3/42 (60) | 1/48 (60) |
| | | 0.0134 $ | NC | 0.2474 | 0.0977 | 0.5000 |
| | Hemangioma | 0/48 (60) | 1/51 (60) | 0/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.7447 | 0.5152 | NC | NC | NC |
| | Trichoepithelioma | 0/48 (60) | 0/50 (60) | 1/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.4759 | NC | 0.5000 | NC | NC |
| I.S. Dors.Tho. Rt | Fibroadenoma | 0/48 (60) | 1/50 (60) | 1/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.5942 | 0.5102 | 0.5000 | NC | NC |
| | Keratoacanthoma | 0/48 (60) | 0/50 (60) | 0/48 (61) | 0/41 (60) | 1/48 (60) |
| | | NC | NC | NC | NC | 0.5000 |
| I.S. Lumbar, Left | Keratoacanthoma | 0/48 (60) | 0/50 (60) | 0/48 (61) | 0/41 (60) | 1/48 (60) |
| | | NC | NC | NC | NC | 0.5000 |
| Kidney | Adenoma: Tubular Cell | 0/48 (60) | 1/50 (60) | 0/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.7433 | 0.5102 | NC | NC | NC |
| | Carcinoma: Tubular Cell | 1/48 (60) | 0/50 (60) | 0/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| | Adenoma: Tubular Cell/ Carcinoma: Tubular Cell | 1/48 (60) | 1/50 (60) | 0/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.9351 | 0.7627 | 1.0000 | 1.0000 | 1.0000 |
| Liver | Adenoma: Hepatocellular | 3/48 (60) | 3/50 (60) | 1/48 (61) | 0/41 (60) | 3/48 (60) |
| | | 0.9777 | 0.6806 | 0.9414 | 1.0000 | NC |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

**Table 2B: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Female Rats (Continued)**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 0 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| Lung | Adenoma: Alveolar/Bronchiolar | 0/48 (60) | 1/50 (60) | 0/48 (61) | 1/42 (60) | 1/48 (60) |
| | | 0.2831 | 0.5102 | NC | 0.4667 | 0.5000 |
| | Carcinoma:Alveolar/Bronchiolar | 0/48 (60) | 0/50 (60) | 1/48 (61) | 1/41 (60) | 0/48 (60) |
| | | 0.1603 | NC | 0.5000 | 0.4607 | NC |
| | Adenoma: Alveolar/Bronchiolar/ Carcinoma:Alveolar/Bronchiolar | 0/48 (60) | 1/50 (60) | 1/48 (61) | 2/42 (60) | 1/48 (60) |
| | | 0.0957 | 0.5102 | 0.5000 | 0.2150 | 0.5000 |
| Mammary Gland | Adenocarcinoma | 3/48 (60) | 2/50 (59) | 2/47 (58) | 1/41 (58) | 2/48 (60) |
| | | 0.7918 | 0.8315 | 0.8126 | 0.9203 | 0.8192 |
| | Adenoma | 1/48 (60) | 0/49 (59) | 1/47 (58) | 1/41 (58) | 2/48 (60) |
| | | 0.3623 | 1.0000 | 0.7474 | 0.7120 | 0.5000 |
| | Adenocarcinoma/Adenoma | 4/48 (60) | 2/50 (59) | 3/47 (58) | 2/41 (58) | 4/48 (60) |
| | | 0.6646 | 0.9070 | 0.7736 | 0.8581 | NC |
| | Fibroadenoma | 17/51 (60) | 22/51 (59) | 20/48 (58) | 15/42 (58) | 11/49 (60) |
| | | 0.4827 | 0.2076 | 0.2583 | 0.4906 | 0.9247 |
| Ovary | Benign Granulosa-Theca Cell Tumor | 0/48 (60) | 0/50 (60) | 1/48 (61) | 1/42 (60) | 0/48 (60) |
| | | 0.1637 | NC | 0.5000 | 0.4667 | NC |
| Pancreas | Adenoma: Islet Cell | 0/48 (60) | 0/50 (60) | 1/48 (61) | 0/41 (60) | 1/48 (60) |
| | | 0.4759 | NC | 0.5000 | NC | 0.5000 |
| Pituitary | Adenoma: Pars Distalis | 37/54 (58) | 32/52 (59) | 31/52 (60) | 28/48 (59) | 33/52 (60) |
| | | 0.8413 | 0.8308 | 0.8766 | 0.8987 | 0.7747 |
| Skin | Adenoma: Basal Cell | 0/48 (60) | 0/50 (60) | 0/48 (61) | 1/42 (60) | 0/48 (60) |
| | | 0.2234 | NC | NC | 0.4667 | NC |
| Skin Miscellaneous | Adenoma: Basal Cell | 0/15 (17) | 0/19 (22) | 0/23 (28) | 0/20 (28) | 1/25 (31) |
| | | NC | NC | NC | NC | 0.6250 |
| | Adenoma: Sebaceous | 0/15 (17) | 0/19 (22) | 1/23 (28) | 0/20 (28) | 0/25 (31) |
| | | 0.5584 | NC | 0.6053 | NC | NC |
| | Keratoacanthoma | 0/15 (17) | 1/19 (22) | 1/23 (28) | 0/20 (28) | 0/25 (31) |
| | | 0.6904 | 0.5588 | 0.6053 | NC | NC |
| | Papilloma: Squamous Cell | 0/15 (17) | 0/19 (22) | 0/23 (28) | 2/21 (28) | 0/25 (31) |
| | | 0.0699 | NC | NC | 0.3333 | NC |
| Spleen | Fibrosarcoma | 0/48 (60) | 1/50 (60) | 0/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.7433 | 0.5102 | NC | NC | NC |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

**Table 2B: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Female Rats (Continued)**

| Organ name | Tumor name | Vehicle (VC) 0 µg P - Trend | Low (L) 10 µg P - VC vs. L | Mid (M) 25 µg P - VC vs. M | High (H) 50 µg P - VC vs. H | Positive (PC) 0 µg P - VC vs. PC |
|---|---|---|---|---|---|---|
| Thyroid | Adenoma: C-Cell | 3/49 (60) | 5/51 (60) | 7/50 (61) | 5/42 (60) | 6/48 (60) |
| | | 0.1679 | 0.3803 | 0.1672 | 0.2739 | 0.2329 |
| | Carcinoma: C-Cell | 1/48 (60) | 1/50 (60) | 1/49 (61) | 0/41 (60) | 1/48 (60) |
| | | 0.8097 | 0.7627 | 0.7577 | 1.0000 | NC |
| | Adenoma: C-Cell/ Carcinoma: C-Cell | 4/49 (60) | 6/51 (60) | 8/50 (61) | 5/42 (60) | 7/48 (60) |
| | | 0.2694 | 0.3963 | 0.1882 | 0.4017 | 0.2500 |
| | Adenoma: Follicular Cell | 0/48 (60) | 1/50 (60) | 1/48 (61) | 1/42 (60) | 3/48 (60) |
| | | 0.2618 | 0.5102 | 0.5000 | 0.4667 | 0.1211 |
| | Carcinoma: Follicular Cell | 1/48 (60) | 1/50 (60) | 1/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.8077 | 0.7627 | NC | 1.0000 | 1.0000 |
| | Adenoma: Follicular Cell/ Carcinoma: Follicular Cel | 1/48 (60) | 2/50 (60) | 2/48 (61) | 1/42 (60) | 3/48 (60) |
| | | 0.5151 | 0.5155 | 0.5000 | 0.7184 | 0.3085 |
| Urinary Bladder | Carcinoma: Transitional Cell | 1/45 (57) | 0/50 (60) | 0/48 (61) | 0/40 (58) | 0/48 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| Uterus | Deciduoma | 0/48 (60) | 1/50 (60) | 1/48 (61) | 1/42 (60) | 0/48 (60) |
| | | 0.2618 | 0.5102 | 0.5000 | 0.4667 | NC |
| | Hemangioma | 0/48 (60) | 1/50 (60) | 0/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.7433 | 0.5102 | NC | NC | NC |
| | Leiomyoma | 0/48 (60) | 0/50 (60) | 0/48 (61) | 1/42 (60) | 0/48 (60) |
| | | 0.2234 | NC | NC | 0.4667 | NC |
| | Polyp: Endometrial Stromal | 17/49 (60) | 10/51 (60) | 13/50 (61) | 12/44 (60) | 12/49 (60) |
| | | 0.6354 | 0.9733 | 0.8770 | 0.8402 | 0.9081 |
| | Sarcoma: Endometrial Stromal | 3/50 (60) | 1/50 (60) | 0/48 (61) | 1/41 (60) | 1/48 (60) |
| | | 0.8509 | 0.9413 | 1.0000 | 0.9138 | 0.9362 |
| | Polyp: Endometrial Stromal/ Sarcoma: Endometrial Stromal | 20/51 (60) | 11/51 (60) | 13/50 (61) | 13/44 (60) | 13/49 (60) |
| | | 0.7285 | 0.9848 | 0.9486 | 0.8858 | 0.9412 |
| Vagina | Benign Granular Cell Tumor | 1/48 (60) | 0/50 (60) | 1/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.7267 | 1.0000 | NC | 1.0000 | 1.0000 |
| | Hemangiosarcoma | 0/48 (60) | 0/50 (60) | 1/49 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.4787 | NC | 0.5052 | NC | NC |
| | Papilloma: Squamous Cell | 1/48 (60) | 0/50 (60) | 0/48 (61) | 0/41 (60) | 1/48 (60) |
| | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | NC |
| | Polyp | 0/48 (60) | 1/50 (60) | 0/48 (61) | 0/41 (60) | 0/48 (60) |
| | | 0.7433 | 0.5102 | NC | NC | NC |
| Whole body | Osteoblastoma | 0/48 (60) | 8/51 (60) | 7/49 (61) | 9/44 (60) | 4/48 (60) |
| | | 0.0064 $ | 0.0037 $ | 0.0067 $ | 0.0008 $ | 0.0586 |
| | Osteosarcoma | 1/48 (60) | 11/52 (60) | 22/53 (61) | 37/55 (60) | 24/52 (60) |
| | | 0.0000# | 0.0030# | 0.0000# | 0.0000# | 0.0000# |
| | Osteosarcoma/Osteoblastoma | 1/48 (60) | 16/52 (60) | 27/53 (61) | 40/55 (60) | 27/52 (60) |
| | | 0.0000# | 0.0001# | 0.0000# | 0.0000# | 0.0000# |

& X/YY (ZZ): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

**Figure 1A: Kaplan-Meier Survival Functions for Male Rats**

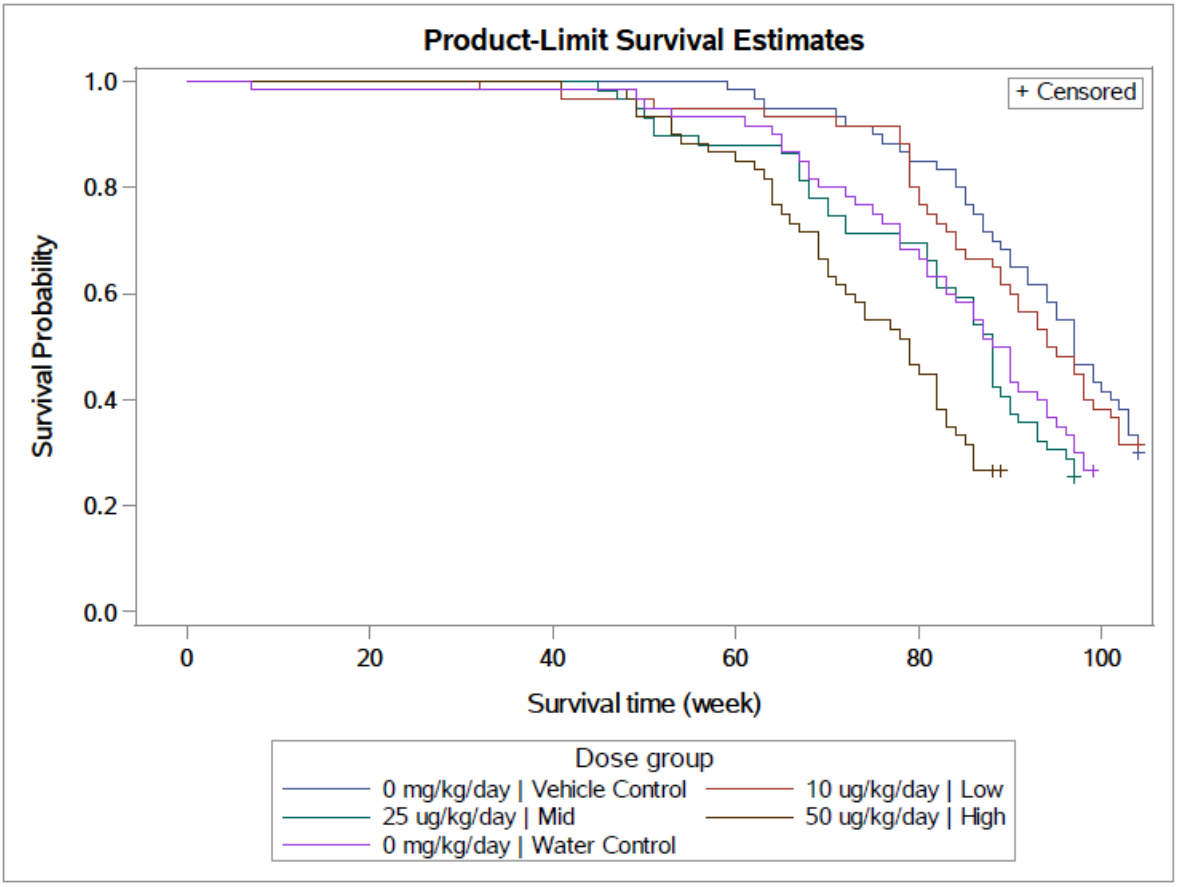
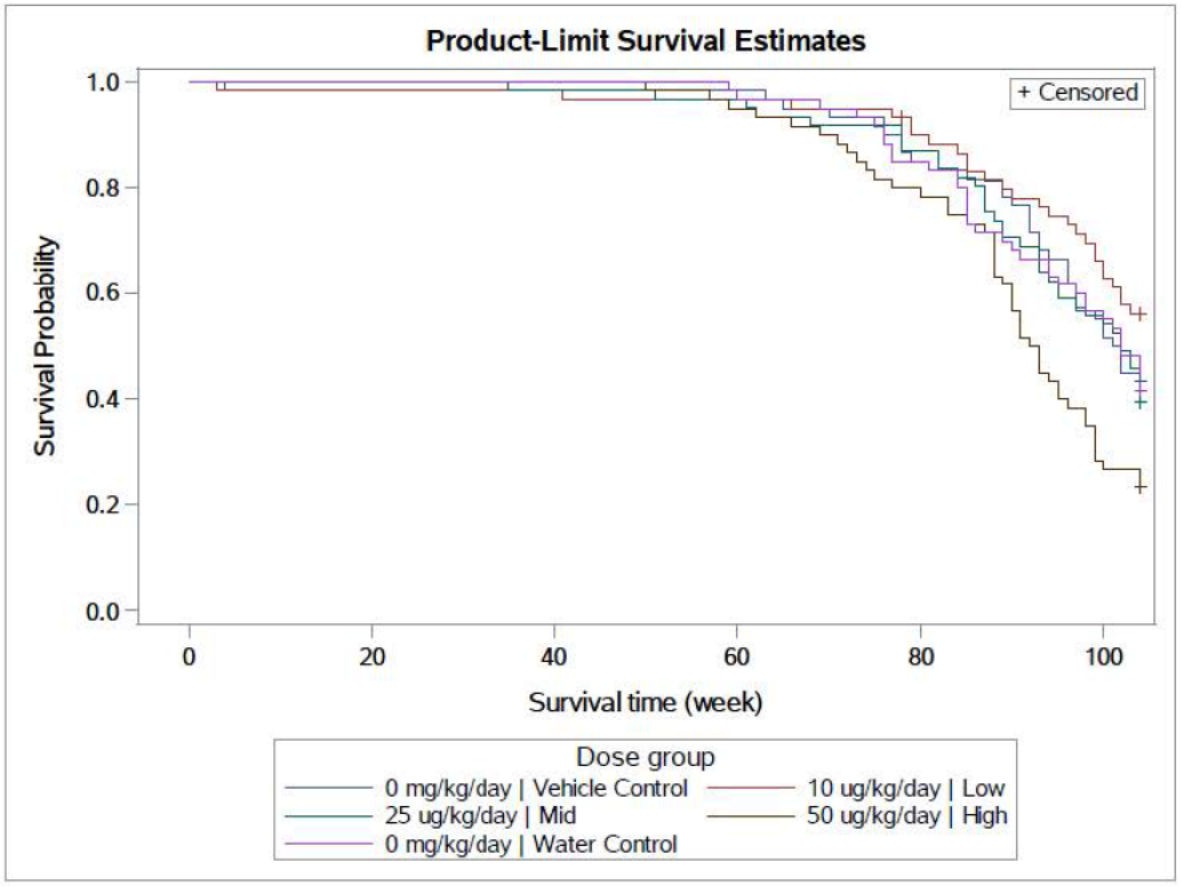


**Figure 1B: Kaplan-Meier Survival Functions for Female Rats**



## 5. References

Bailer, A.J, Portier, C.J. (1988). “Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples.” *Biometrics*, 44, 417-431.

Bieler, G.S. and Williams, R.L. (1993). “Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity”. *Biometrics* 49, 793-801.

Guidance for Industry. Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Statues of Pharmaceuticals (Draft Guidance). U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). May 2001.

Haseman, J. (1983). “A re-examination of false-positive rates for carcinogenesis studies”, *Fundamental and Applied Toxicology*, 3: 334-339.

Lin K.K. and Rahman A.M. (1998)” Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs”, *Journal of Biopharmaceutical Statistics*, 8(1), 1-15.

Peto, R., M.C. Pike, N.E. Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, Richards, and J.Wahrendorf. (1980) “Guidelines for sample sensitive significance test for carcinogenic effects in long-term animal experiments”, Long term and short term screening assays for carcinogens: A critical appraisal, International agency for research against cancer monographs, *Annex to supplement, World Health Organization, Geneva*, 311-426.

Rahman, A.M., and Lin, K.K. (2008), "A Comparison of False Positive Rates of Peto and Poly-3 Methods for Long-Term Carcinogenicity Data Analysis Using Multiple Comparison Adjustment Method Suggested by Lin and Rahman", *Journal of Biopharmaceutical Statistics*, 18:5, 849-858.

Rahman, A.M., and Lin, K.K. (2009), "Design and Analysis of Chronic Carcinogenicity Studies of Pharmaceuticals in Rodents", in “Design and Analysis of Clinical Trials with Time-to-Event Endpoints", K.E Peace, Editor, Chapman & Hall/CRC, Taylor & Francis Group, LLC, Boca Raton, FL, London, and New York.

---

**This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.**

---

/s/

---

HEPEI CHEN
12/14/2016

KARL K LIN
12/14/2016
Concur with review

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

# STATISTICAL REVIEW AND EVALUATION

## CLINICAL STUDIES

| | |
|---|---|
| **NDA/ Serial Number:** | 208743 / 0001 |
| **Drug Name:** | Abaloparatide subcutaneous injection |
| **Indication:** | Treatment of postmenopausal women with osteoporosis |
| **Applicant:** | Radius Health |
| **Date(s):** | Submission date: 3/30/2016 |
| | PDUFA date: 3/30/2017 |
| **Review Priority:** | Standard |
| **Biometrics Division:** | Division of Biometrics III |
| **Statistical Reviewer:** | Jia Guo, Ph.D. |
| **Concurring Reviewer:** | Sonia Castillo, Ph.D. (acting team leader) |
| **Medical Division:** | Division of Bone, Reproductive and Urologic Products |
| **Clinical Team:** | Stephen Voss , M.D., Clinical Reviewer |
| | Theresa Kehoe, M.D., Clinical Team Leader |
| **Project Manager:** | Samantha Bell |

**Keywords: incidence rate, risk reduction, relative risk, multiplicity, ANCOVA**

# Table of Contents

# LIST OF TABLES



# LIST OF FIGURES

# 1 EXECUTIVE SUMMARY

The Applicant submitted one randomized, double-blind phase 3 study and the first 6 months of data for its extension study to demonstrate the efficacy of abaloparatide in women with postmenopausal osteoporosis at high risk for fracture.

The analysis results show that abaloparatide:

- Reduced the incidence rate of new vertebral fracture compared to placebo (0.55% vs. 4.2%, P-value<0.0001) at month 18. The relative risk reduction in new vertebral fracture was 86%.
- Increased bone mineral density at the total hip, femoral neck and lumbar spine by 3.5%, 3.3% and 8.7% compared to placebo at month 18.
- Reduced the non-vertebral fracture event rate at month 19 compared to placebo (2.7% vs. 4.7%, P-value=0.045).

At month 25:

- There was a reduction in the incidence of new vertebral fracture in the former abaloparatide group compared to the former placebo group from baseline to month 25 (0.55% versus 4.40%, respectively, p<0.0001). There was an 87% relative risk reduction in new vertebral fracture.
- There was a continued increase in bone mineral density at the total hip, femoral neck and lumbar spine in the former abaloparatide group compared to the former placebo (4.1%, 4.1% and 9.3%).
- There was a continued benefit in nonvertebral fracture with a 52% reduction in risk for the former abaloparatide group compared to the former placebo group (2.7% versus 5.6%; p = 0.0168).

The two studies provide evidence demonstrating the efficacy of abaloparatide for the treatment of postmenopausal osteoporosis in women at high risk of fracture.

However, the submitted studies enrolled very few U.S. patients and the efficacy of abaloparatide was not evaluable in the U.S. patients from a statistical perspective. The issue of generalizability of the efficacy results in non-U.S. regions to the U.S. population will be a clinical decision.

# 2 INTRODUCTION

## 2.1 Overview

The Applicant, Radius Health, seeks approval of abaloparatide subcutaneous injection for the treatment of postmenopausal osteoporosis in women. The proposed dose is 80 μg once daily administered by subcutaneous self-injection to the periumbilical region.

According to the Applicant:

> Abaloparatide is a novel, synthetic 34-amino acid peptide analog of human PTH-related peptide (hPTHrP), with molecular modifications of specific amino acids. Abaloparatide has enhanced PTH1 Receptor RG/RO selectivity as compare to PTH or PTHrP and has demonstrated retention of the potent anabolic activity of PTH, with reduced bone resorption and calcium-mobilizing potential. Thus, abaloparatide is expected to have similar or greater efficacy in restoring BMD in individuals with osteoporosis than hPTH(1-34), but with less risk of causing hypercalcemia.

This statistical review is based on Phase 3 Study BA-58-05-003 along with the first 6 months of data from its 2-year extension Study BA-058-05-005. Table 1 presents a brief summary of both studies.

**Table 1: List of All Studies Included in the Statistical Review**

| Study | Phase and Design | Treatment Period | # of Randomized Subjects per Arm | Study Population |
|---|---|---|---|---|
| BA-58-05-003 | Phase 3, randomized, double-blind, placebo-controlled | Screening: 2 months<br>Pre-treatment: 1 week<br>Treatment: 18 months<br>Follow-uo:1 month | Abaloparatide: 824<br>Teriparatide: 818<br>Placebo: 821 | • Healthy ambulatory postmenopausal women from 50 to 85 years of age with osteoporosis.<br>• Have a BMD T score ≤2.5 and >-5.0 at the lumbar spine or hip (femoral neck) and radiological evidence of 2 or more mild or one or more moderate lumbar or thoracic vertebral fractures, or history of low trauma forearm, humerus, sacral, pelvic, hip, femoral, or tibial fracture within the past 5 years.<br>• Postmenopausal women older than 65 who meet the above fracture criteria but had a T-score ≤2.0 and >-5.0 could be enrolled. Women older than 65 who did not meet the fracture criteria could be enrolled if their T-score is ≤3.0 and >-5.0. |
| BA-58-05-005 | Open-label Extension | 24 months | All patients received Alendronate | Patients who previously received abaloparatide-SC or placebo in Study BA-58-05-003. |

Source: Reviewer's summary.

### 2.2 Data Sources

The study reports, data and additional information were submitted electronically. These items are located in the Electronic Document Room at \\Cdsesub1\evsprod\NDA208743 under submission dates 03/30/2016, 08/15/2016, and 11/22/2016.

## 3 STATISTICAL EVALUATION

### 3.1 Data and Analysis Quality

The Applicant submitted legacy data, analysis data and statistical programs for both studies. Data sets were complete and documented. Statistical analyses of efficacy endpoints in each study were carried out following the pre-specified statistical analysis plan.

### 3.2 Evaluation of Efficacy

This statistical review is based on Phase 3 Study BA-58-05-003 (subsequently referred to as Study 003) along with the first 6 months of data from its 2-year extension Study BA-058-05-005 (subsequently referred to as Study 005).

#### 3.2.1 Study Design and Endpoints

**Study 003**

Study 003 was a randomized, double-blind, placebo-controlled, phase 3 multicenter, international study of the safety and efficacy of abaloparatide subcutaneous injection in ambulatory, postmenopausal (≥ 5 years) women aged 50 to 85 years with severe osteoporosis and at risk of fracture. Main inclusion criteria included:

- a BMD T-score ≤ -2.5 and > -5.0 at the lumbar spine or hip (femoral neck) and radiological evidence of 2 or more mild or one or more moderate lumbar or thoracic vertebral fractures, or
- a history of low trauma forearm, humerus, sacral, pelvic, hip, femoral, or tibial fracture within the past 5 years, or
- women older than 65 who meet the fracture criteria but had a T-score ≤ -2.0 and > -5.0, or
- women older than 65 who do not meet the fracture criteria could also be enrolled if their T-score was ≤ -3.0 and > -5.0.

After screening, eligible subjects were equally randomized to either abaloparatide (80 μg), teriparatide (20 μg), or matching placebo, each self-administered as a daily subcutaneous injection. Calcium and vitamin D supplements were taken daily in the evening.

Teriparatide was included in this study to evaluate a safety endpoint of interest, hypercalcemia, and any safety or efficacy data related to it will not be further discussed in this review.

The study consisted of a screening period (up to 2 months), a pretreatment period (1 week), the treatment period (18 months) and a follow-up period (1 month). The total duration of study participation was approximately 20 to 21 months.

Efficacy assessments included evaluation by x-ray after 18 months of treatment (end of treatment) and evaluations of bone mineral density (BMD) by dual energy x-ray absorptiometry (DXA) after 6, 12, and 18 months of treatment. Serum markers of bone metabolism, abaloparatide antibody, and abaloparatide serum levels were also measured during the treatment period.

The primary study objective was to evaluate abaloparatide compared to placebo with respect to new vertebral fractures.

The key secondary objectives were to evaluate abaloparatide compared to placebo with respect to:
- BMD of the lumbar spine, hip, and femoral neck
- nonvertebral fractures

The primary efficacy endpoint was the percentage of patients with one or more incidents of x-ray verified new vertebral fracture from baseline to month 18.

The key secondary efficacy endpoints were:
- percent change from baseline to month 18 in lumbar spine BMD
- percent change from baseline to month 18 in total hip BMD
- percent change from baseline to month 18 in femoral neck BMD
- time to first incident nonvertebral fracture (NVF) by the follow-up visit (month 19)

The two efficacy populations used for analyses are described below:
- The Intent-to-Treat (ITT) population includes all patients who were randomized into the study and received the randomized study medication kit on Day 1.
- The modified Intent-to-Treat (mITT) population includes all ITT patients who have both the pre-treatment and the post-baseline evaluable radiologic assessment (spine X-ray).

**Multiplicity Control**
To control the overall type I error, the primary and secondary efficacy endpoints were tested in the sequential order presented below. Each of the tests was performed comparing abaloparatide to placebo at the 2-sided 5% significance level. If at any step of the sequential testing treatment difference was not statistically significant, then all subsequent comparisons cannot be claimed to be statistically significant.

1. Incidence of vertebral fracture
2. Total hip BMD
3. Femoral neck BMD
4. Lumbar spine BMD
5. Nonvertebral fracture

**Study 005**

Study 005 was a 24-month extension open label study of patients who completed 18 months of abaloparatide or placebo treatment in Study 003. All patients enrolled in study 005 took alendronate treatment (70 mg orally once per week). Patients and investigators who participated in Study 005 remained blinded to prior treatment assignment as part of Study 003 through the first 6 months of Study 005 (Visit 3 in 005).

The first 6 months data from Study 005 was used to confirm the efficacy of abaloparatide for prevention of new vertebral fracture up to month 25 in patients who had completed Study 003 and received 18 months of blinded treatment with abaloparatide or placebo and then transitioned to 6 months of treatment with alendronate.

The primary efficacy endpoint was the percentage of patients with one or more incidents of x-ray verified new vertebral fracture from baseline (Visit 1 in Study 003) to month 25 (Visit 3 in Study 005).

The following key secondary efficacy endpoints comparing abaloparatide to placebo were tested in the following sequential order to control type I error in the same manner as previously described:

- percent change from baseline to month 25 in total hip BMD
- percent change from baseline to month 25 in femoral neck BMD
- percent change from baseline to month 25 in lumbar spine BMD
- time to first incident nonvertebral fracture (NVF) by month 25

The two efficacy populations used for analyses are described below:

- The Intent-to-Treat (ITT) population includes all Study 003 patients who enrolled into Study 005.
- The modified Intent-to-Treat (mITT) population includes all Study 003 ITT patients who have both the pre-treatment and the post-baseline month 25 (Visit 3 in Study 005) evaluable radiologic assessment (spine X-ray).

### 3.2.2 Statistical Methodologies

**Study 003**

The primary efficacy endpoint, percentage of patients with one or more incidents of new vertebral fracture, was analyzed using the Wilson's score method to obtain the estimate and the 95% confidence interval for each treatment group using the mITT population. The absolute risk reduction and its corresponding 95% confidence interval for the treatment difference were derived using the Newcombe approach. The Fisher's exact test was used to compare the two treatment groups with respect to the primary efficacy endpoint.

The primary population for the analysis of BMD data was the ITT population. An analysis of covariance (ANCOVA) model was used to compare treatment groups for the percent change from baseline in BMD with missing imputation based on last observation (on treatment) carried forward (LOCF). The treatment comparison was derived by testing the contrast (difference in least squares mean) between the two treatment groups at each scheduled study visit (month 6, month 12 and month 18).

The percent change from baseline in BMD was the dependent variable for the analysis of covariance (ANCOVA) model with LOCF and MMRM (as a sensitivity analysis). Each model included fixed effects (DXA instrument manufacturer [Hologic vs Lunar Prodigy], treatment, visit, and treatment-by-visit interaction) and fixed covariate

(baseline BMD). For each treatment comparison, the ANCOVA model with LOCF was constructed and generated using the data from the two treatment groups that were to be compared.

For the secondary efficacy endpoint, nonvertebral fractures, the log-rank test was used to compare the difference in time to first nonvertebral fracture between the abaloparatide and placebo groups using the ITT population. The Cox proportional hazard model was used to calculate the hazard ratio (95% CI) of incident nonvertebral fractures between the two treatment groups. The Kaplan-Meier (K-M) curve was generated to graphically display the data and the incidence rates were estimated using the K-M method at 19 months.

Other secondary efficacy endpoints were analyzed for exploratory purposes and are not presented here.

**Study 005**
The Study 003 and Study 005 integrated analysis was performed using the double-blind treatment group (placebo or abaloparatide) that was assigned in Study 003. Efficacy analyses up to month 25 (following the first 6-month treatment period with alendronate [Visit 3 in 005]) were performed.

The primary efficacy endpoint comparison was performed using the Study 005 mITT population. The percentage of patients who had one or more new vertebral fractures and 95% confidence interval was provided for each study 003 double-blind treatment group and overall using Wilson's score method. The Fisher's exact test was used to compare the two Study 003 double-blind treatment groups (abaloparatide and placebo) with respect to the primary efficacy endpoint. The absolute risk reduction and its 95% confidence interval for the treatment difference were derived using the Newcombe method (1998). The relative risk reduction and its 95% confidence interval were produced using the Wald method.

The BMD secondary efficacy endpoints were analyzed using the Study 005 ITT population at month 25 (Visit 3 in Study 005) compared to baseline (Visit 1 in Study 003). An ANCOVA model was used to compare treatment groups for the percent change from baseline in BMD with missing data imputation based on LOCF (last observation carried forward). If a value was missing at month 25, it was imputed using the value from month 18 (Visit 9 in Study 003).

The percent change from baseline in BMD was the dependent variable for the ANCOVA model, which included fixed effects (DXA instrument manufacturer [Hologic vs Lunar Prodigy], treatment) and a fixed covariate (baseline BMD [Visit 1 in Study 003]).

For the secondary efficacy endpoint, nonvertebral fracture (NVF), the statistical analysis utilized (a) the log-rank test for inferential statistics and (b) the Kaplan-Meier method for estimates of event rates. The log-rank test was the primary analysis method used to compare the difference in time to first NVF between treatment groups up to month 25. The Cox proportional hazard model, including treatment group as fixed effect, was used to calculate the hazard ratio (95% CI) of incident NVF between the two treatment groups. A Kaplan-Meier curve was generated to graphically display the data and the incidence rates were estimated using the Kaplan-Meier method at month 25.

Other secondary efficacy endpoints were analyzed for exploratory purposes and are not presented here.

### 3.2.3 Patient Disposition, Demographic and Baseline Characteristics

Patient disposition for both studies is summarized in Table 2 and Table 3. In Study 003, a total of 2463 patients were randomized to three treatment groups. Of these, a total of 1645 (69.2%) patients were randomized to the placebo and abaloparatide groups. The discontinuation rate in Study 003 was 22.4% in the placebo group and 26.5% in the abaloparatide group. The most common reason for discontinuation in both studies was adverse events.

**Table 2: Study 003 - Summary of Patient Disposition**

| | Placebo n (%) | Abaloparatide n (%) | Overall N (%) |
|---|---|---|---|
| Overall Randomized | 821 (100.0) | 824 (100.0) | 1645 (100.0) |
| Completed Study | 637 (77.6) | 606 (73.5) | 1243 (75.6) |
| Discontinued Study | 184 (22.4) | 218 (26.5) | 402 (24.4) |
| Reason for Discontinuation [1] | | | |
| Adverse Event | 53 (28.8) | 89 (40.8) | 142 (35.3) |
| Withdrew Consent | 48 (26.1) | 47 (21.6) | 95 (23.6) |
| Refusal of Treatment | 33 (17.9) | 31 (14.2) | 64 (15.9) |
| Lost to follow up | 5 (2.7) | 15 (6.9) | 20 (5.0) |
| Inability to Complete Study Procedures | 7 (3.8) | 11 (5.0) | 18 (4.5) |
| Non-compliance | 10 (5.4) | 6 (2.8) | 16 (4.0) |
| Serious Intercurrent Illness | 0 | 4 (1.8) | 4 (1.0) |
| Protocol Violation | 4 (2.2) | 4 (1.8) | 8 (2.0) |
| Patient died during the study | 5 (2.7) | 3 (1.4) | 8 (2.0) |
| Continuing significant deterioration from baseline (>7%) of BMD at lumbar spine or hip (after confirmation of the findings) | 12 (6.5) | 1 (0.5) | 13 (3.2) |
| Administrative Reasons | 0 | 1 (0.5) | 1 (0.2) |
| Hypercalcemia or Hypercalciuria | 0 | 1 (0.5) | 1 (0.2) |
| Treatment related SAE | 0 | 0 | 0 (0.0) |
| Hypersensitivity to abaloparatide/placebo/teriparatide | 1 (0.5) | 0 | 1 (0.2) |
| Other | 6 (3.3) | 5 (2.3) | 11 (2.7) |

[1] Percentages based on the number of patients who did not complete Study 003.
Source: Reviewer' analysis and Table 6 in Study 003 study report.

**Table 3: Study 005 - Summary of Patient Disposition**

| | Placebo n (%) | Abaloparatide n (%) | Overall N (%) |
|---|---|---|---|
| Study 003 Randomized | 821 (100.0) | 824 (100.0) | 1645 (100.0) |
| Entered Study 005 [1] | 581 (70.9) | 558 (67.9) | 1139 (69.2) |
| Ongoing at 6 months [2] | 543 (93.5) | 529 (94.8) | 1027 (94.1) |
| Completed Study [2] | 1 (0.2) | 0 | 1 (0.1) |
| Discontinued Study[2] | 37 (6.4) | 29 (5.2) | 66 (5.8) |
| Reasons for discontinuation of study 005 [3] | | | |
| Adverse Event | 21 (56.8) | 17 (58.6) | 38(57.6) |
| Withdrew Consent | 7 (18.9) | 7 (24.1) | 14 (21.2) |
| Refusal of treatment | 2 (5.4) | 3 (10.3) | 5 (7.6) |
| Inability to complete study procedure | 2 (5.4) | 0 | 2 (3.0) |
| Lost to follow up | 2 (5.4) | 1 (3.4) | 3 (4.5) |
| Patient died during the study | 1 (2.7) | 1 (3.4) | 2 (3.0) |
| Hypersensitivity to alendronate | 1 (2.7) | 0 | 1(1.5) |
| Other | 1 (2.7) | 0 | 1 (1.5) |

[1] Percentages based on Study 003 number of patients randomized.
[2] Percentages based on number of patients who entered Study 005.
[3] Percentages based on the number of patients who did not complete Study 005.
Source: Reviewer's analysis and Table 4 in Study 005 study report.

The number of patients in each analysis population is presented in Table 4 and Table 5.

**Table 4: Study 003 - Summary of Efficacy Analysis Sets**

| Analysis population | Placebo n (%) | Abaloparatide n (%) | Overall N (%) |
|---|---|---|---|
| Overall Randomized | 821 (100.0) | 824 (100.0) | 1645 (100.0) |
| ITT Population | 821 (100.0) | 824 (100.0) | 1645 (100.0) |
| mITT Population | 711 (86.6) | 690 (83.7) | 1401 (85.2) |

Percentages are based number of randomized patients.
Source: Reviewer's analysis and Table 6 in Study 003 report.

**Table 5: Study 005 - Summary of Efficacy Analysis Sets**

| Analysis population | Placebo n (%) | Abaloparatide n (%) | Overall N (%) |
|---|---|---|---|
| Entered Study 005 | 581 (100.0) | 558 (100.0) | 1139 (100.0) |
| ITT Population | 581 (100.0) | 558 (100.0) | 1139 (100.0) |
| mITT Population | 568 (97.8) | 544 (97.5) | 1112 (97.6) |

Percentages are based number of patients who entered Study 005.
Source: Reviewer's analysis and Table 4 in Study 005 report.

Demographic and baseline characteristics are summarized in the Appendix (Table 18 and Table 19) for each of the two studies. In Study 003, demographic and baseline characteristics were balanced among treatment groups. All patients were postmenopausal women aged 49 to 86 years, inclusive. The median age was 68 years (mean [SD] 68.8 [6.52]) with 19% of women being ≥75 years of age. The median number of years since menopause was 20. Approximately 80% of the patients were white. The mean body mass index (BMI) was approximately 25 $kg/m^2$. The majority of patients were from Europe (55.6%) and South America (27.1%). Only 39 patients were from the U.S. (1.1%). The demographic and baseline characteristics are similar in both studies. The treatment groups were well balanced with respect to BMD T-scores and prevalent vertebral fractures at baseline (see Appendix).

### 3.2.4 Results and Conclusions

This reviewer was able to replicate the Applicant's analysis results.

#### 3.2.4.1 Analysis results for Study 003

***New Vertebral Fracture***

Table 6 presents the primary efficacy results for Study 003. The incidence of new vertebral fracture was 0.58% in the abaloparatide group compared to 4.22% in the placebo group (P-value: <0.0001). The relative risk reduction in new vertebral fracture was 86% in patients receiving abaloparatide compared to placebo. Of note is that no new vertebral fractures occurred in U.S. patients.

**Table 6: Study 003 - Incidence of New Vertebral Fracture (mITT population)**

| Incidence Parameter | Abaloparatide (N=690) | Placebo (N=711) |
|---|---|---|
| n/N (%)<br>95% CI[1]<br>P-value[2] | 4/690 (0.58%)<br>(0.23%, 1.48%)<br><0.0001 | 30/711 (4.22%)<br>(2.97%, 5.96%) |
| Risk reduction vs. placebo[3]<br>95% CI | -3.64%<br>(-5.42%, -2.10%) | |
| Relative risk reduction vs. placebo[4]<br>95% CI | -86%<br>(-95%, -61%) | |

Source: Reviewer's analysis and Table 13 in Study 003 report.
[1] 95% CI for percentage was based on the Wilson's Score method.
[2] P-value from Fisher's exact test comparing abaloparatide-SC with placebo.
[3]The risk reduction was calculated as (Abaloparatide-SC-Placebo). 95% CI was based on Newcombe's method.
[4] The RRR was calculated as (Abaloparatide - Placebo)/Placebo. 95% CI was based on Wald's method

***Bone Mineral Density***

Baseline bone mineral density (BMD) values were similar across all treatment groups for the total hip; femoral neck and lumbar spine. According to the pre-specified fixed-sequence testing procedure, the BMD analysis results (Table 7) showed that abaloparatide increased BMD from baseline to 18 months at the total hip, femoral neck and lumbar spine ($p<0.0001$ for all three sites) compared to placebo.

**Table 7: Study 003 - Summary of Bone Mineral Density at Month 18 (ITT population, LOCF)**

| Location | Statistic | Placebo (N = 821) | Abaloparatide (N = 824) |
|---|---|---|---|
| Total Hip | Baseline (SD)<br>% Change from baseline (SD) | 0.77 (0.10)<br>-0.1 (2.8) | 0.77 (0.09)<br>3.4 (3.5) |
| | LS mean diff vs. placebo (SE)<br>95% CI<br>P-value vs. placebo | | 3.5 (0.1)<br>(3.3, 3.8)<br><0.0001 |
| Femoral neck | Baseline (SD)<br>% Change from baseline (SD) | 0.73 (0.10)<br>-0.4(3.6) | 0.73 (0.09)<br>2.9 (4.2) |
| | LS mean diff vs. placebo<br>95% CI<br>P-value vs. placebo | | 3.3 (0.2)<br>(3.0, 3.7)<br><0.0001 |
| Lumbar Spine | Baseline (SD)<br>% Change from baseline (SD) | 0.82 (0.10)<br>0.5 (4.0) | 0.83 (0.11)<br>9.2 (7.5) |
| | LS mean diff vs. placebo<br>95% CI<br>P-value vs. placebo | | 8.7 (0.3)<br>(8.2, 9.2)<br><0.0001 |

Source: Reviewer's analysis; Table 14.2.7.1A, Table 14.2.7.2A and Table 14.2.7.3A
LS mean difference vs. placebo, 95% CI and P-values were obtained from ANCOVA model.

***Nonvertebral Fracture***

Table 8 presents the nonvertebral fracture results. Abaloparatide prolonged the time to the first incidence of nonvertebral fracture compared to placebo (log-rank p=0.0448). The Kaplan-Meier event rate at month 19 was lower in the abaloparatide group compared to the placebo group (2.7% versus 4.7%, respectively), with a 43% reduction in the hazard of nonvertebral fracture (hazard ratio 0.57, 95% CI [0.32, 1.00]). The Kaplan-Meier curve was consistently lower for the abaloparatide group compared to the placebo group at any time-point during the overall 19 months of the observational period.

Per the clinical reviewer's request, the incidence rate of non-vertebral fracture was also analyzed using the same approach for the vertebral fractures. The relative risk reduction in non-vertebral fracture was 46% in patients receiving abaloparatide compared to placebo.

**Table 8: Study 003 – Summary of Non-Vertebral Fractures (NVF) (ITT population)**

| | **Placebo (N = 821)** | **Abaloparatide (N = 824)** |
|---|---|---|
| K-M Estimated Event Rate at 19 Months | 4.7% | 2.7% |
| Number of Patients with Event n (%) | 33 (4.0) | 18 (2.2) |
| Number of Patients Censored n (%) | 788 (96.0) | 806 (97.8) |
| Hazard Ratio vs Placebo (95% CI) [1] | - | 0.57 (0.32, 1.00) |
| P-value vs Placebo [2] | - | 0.0489 |
| Absolute Risk Reduction vs. placebo (%) (95% CI) [3] | | -1.84 (-3.60, -0.15) |
| Relative Risk Reduction (%) vs. placebo (95% CI) [4] | | -46 (-69, -4) |

Source: Reviewer's analysis and Table 15 in Study 003 report.

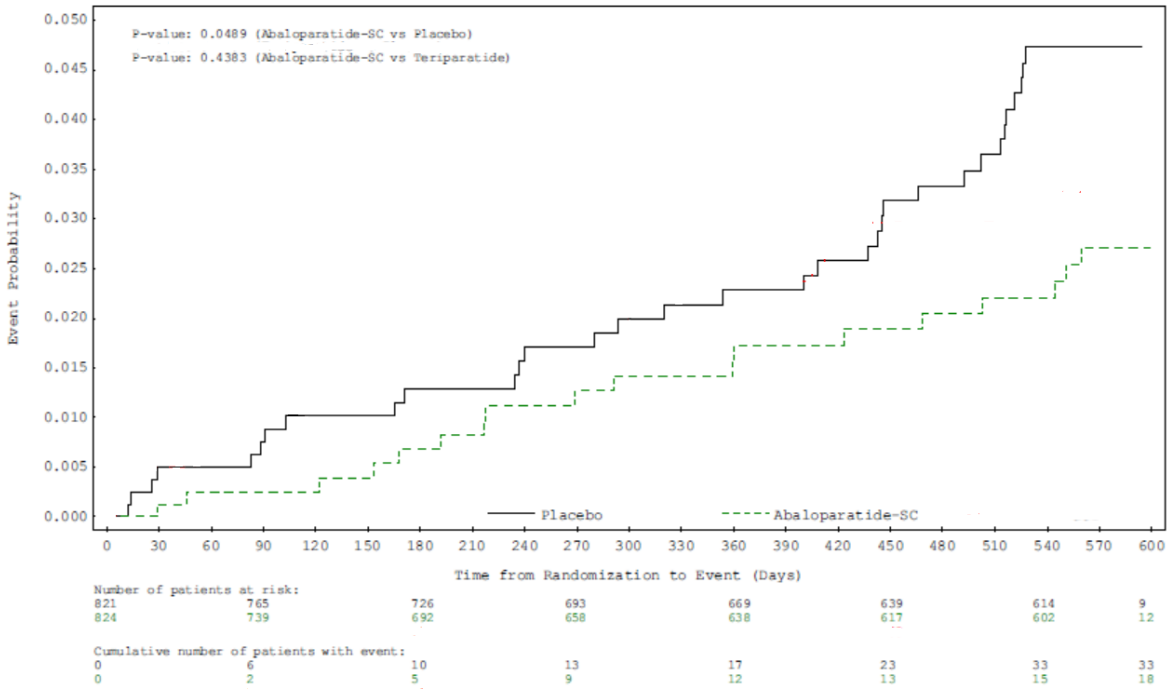
[1] Cox proportional hazard model was used to calculate the hazard ratio with placebo as reference.

[2] P-values were from the log rank test.

[3]The risk reduction was calculated as (Abaloparatide-SC-Placebo). 95% CI was based on Newcombe's method.

[4] The RRR was calculated as (Abaloparatide - Placebo)/Placebo. 95% CI was based on Wald's method

**Figure 1: Kaplan-Meier Curve of Time to First Incidence of Non-Vertebral Fracture by Treatment Group (ITT population)**




Source: Figure 3 in study 003 report.

### 3.2.4.2 Analysis results for Study 005

In Study 005, there were no reports of new vertebral fractures in the former abaloparatide group at month 6, whereas there were 6 patients in the former placebo group that had new vertebral fractures (Table 9). There was a reduction in the incidence of new vertebral fracture in the former abaloparatide group compared to the former placebo group from Study 003 through month 25 (0.55% versus 4.40%, respectively, p<0.0001).

At month 25, there was an 87% relative risk reduction (RRR) in new vertebral fractures in patients previously randomized to abaloparatide compared to placebo in Study 003. Based on the Study 005 mITT population and using only the Study 003 data, there was an 84% RRR in new vertebral fractures at 18 months for patients previously receiving abaloparatide compared to versus placebo (0.55% versus 3.35%, respectively, p<0.0008).

**Table 9: Incidence of First New Vertebral Fracture for Patients Who Were in Both Study 003 and Study 005 (005 mITT population)**

| | Baseline to Month 18 (Study 003) | | Baseline to Month 25 (Study 003 + Study 005) | |
|---|---|---|---|---|
| **Parameter** | **Abaloparatide (N=544)** | **Placebo (N=568)** | **Abaloparatide (N=544)** | **Placebo (N=568)** |
| **n/N (%)**<br>**95% CI[1]**<br>**P-value[2]** | 3/544 (0.55%)<br>(0.19%, 1.61%) | 19/568 (3.35%)<br>(2.15%, 5.17%)<br><0.0001 | 3/544 (0.55%)<br>(0.19%, 1.61%)<br><0.0001 | 25/568 (4.40%)<br>(3.0%, 6.42%) |
| **Risk reduction vs. placebo[3]**<br>**95% CI** | -2.79%<br>(-4.65%, -1.20%) | | -3.85%<br>(-5.90%, -2.09%) | |
| **Relative risk reduction vs. placebo[4]**<br>**95% CI** | -83.5%<br>(-95.0%, -44.6%) | | -87.5%<br>(-96.2%, -59.8%) | |

Source: Table 11 in Study 005 report. Only new vertebral fractures were included in the analysis.
* Treatment groups were based on Study 003 randomization, all patients in extension Study 005 received alendronate
[1] 95% CI for percentage was based on the Wilson's Score method.
[2] P-value from Fisher's exact test comparing abaloparatide-SC with placebo.
[3]The risk reduction was calculated as (Abaloparatide-SC-Placebo). 95% CI was based on Newcombe's method.
[4] The RRR was calculated as (Abaloparatide - Placebo)/Placebo. 95% CI was based on Wald's method

There was one subject in the placebo group who had one nonvertebral fracture in each study time period.

Increases in BMD were observed when previously treated abaloparatide/placebo patients were transitioned to subsequent treatment with alendronate. Based on the ITT population, baseline BMD values (Visit 1 in Study 003) were similar between both treatment groups for total hip, femoral neck and lumbar spine. Increases in BMD were observed in both the former abaloparatide and former placebo groups at month 25 with differences favoring the former abaloparatide group (p<0.0001) at the total hip, femoral neck and lumbar spine.

**Table 10: Study 005 - Summary of BMD at Month 25 (005 ITT population, LOCF)**

| **Location** | **Statistic** | **Placebo (N = 581)** | **Abaloparatide (N = 558)** |
|---|---|---|---|
| **Total Hip** | Baseline (SD)<br>% Change from baseline (SD) | 0.77 (0.10)<br>1.4 (3.0) | 0.77 (0.09)<br>5.4 (4.0) |
| | LS mean diff vs. placebo (SE)<br>95% CI<br>P-value vs. placebo | | 4.1 (0.2)<br>(3.7, 4.5)<br><0.0001 |
| **Femoral neck** | Baseline (SD)<br>% Change from baseline (SD) | 0.73 (0.10)<br>0.5 (3.8) | 0.73 (0.09)<br>4.5 (4.8) |
| | LS mean diff vs. placebo<br>95% CI<br>P-value vs. placebo | | 4.1 (0.2)<br>(3.6, 4.6)<br><0.0001 |
| **Lumbar Spine** | Baseline (SD)<br>% Change from baseline (SD) | 0.83 (0.10)<br>3.5 (4.3) | 0.83 (0.11)<br>12.8 (8.0) |
| | LS mean diff vs. placebo<br>95% CI<br>P-value vs. placebo | | 9.2 (0.3)<br>(8.6, 9.9)<br><0.0001 |

Source: Reviewer's analysis, Tables 14.2.7.1A, 14.2.7.2A, 14.2.7.3A in study 005 report.
LS mean difference vs. placebo, 95% CI and P-values were obtained from ANCOVA model.

At 25 months from Study 003 baseline in Study 005, the non-vertebral benefit continued with a 52% reduction in the risk of non-vertebral fractures in the former abaloparatide group versus former placebo group [K-M Estimates 2.7% versus 5.6%; hazard ratio 0.48, p = 0.0168].

**Table 11: Summary of Nonvertebral Fractures from Baseline through Month 25 (005 ITT population)**

| | Placebo (N = 581) | Abaloparatide (N = 588) |
|---|---|---|
| K-M Estimated Event Rate at 25 Months | 5.6% | 2.7% |
| Number of Patients with Event n (%) | 32 (5.5) | 15 (2.7) |
| Number of Patients Censored n (%) | 549 (94.5) | 543 (97.3) |
| Hazard Ratio vs Placebo (95% CI) [1] | - | 0.48 (0.26, 0.89) |
| P-value vs Placebo [2] | - | 0.0168 |
| Absolute Risk Reduction vs. placebo (%) (95% CI) [3] | | -2.82 (-5.23, -0.50) |
| Relative Risk Reduction (%) vs. placebo (95% CI) [4] | | -51 (-73, -0.1) |

Source: Reviewer's analysis and Table 15 in Study 003 report.

[1] Cox proportional hazard model was used to calculate the hazard ratio with placebo as reference.

[2] P-values were from the log rank test.

[3]The risk reduction was calculated as (Abaloparatide-SC-Placebo). 95% CI was based on Newcombe's method.

[4] The RRR was calculated as (Abaloparatide - Placebo)/Placebo. 95% CI was based on Wald's method

### 3.3 Evaluation of Safety

Refer to the clinical reviewer's review for evaluation of safety data.

## 4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

### 4.1 Race, Age, and Geographic Region

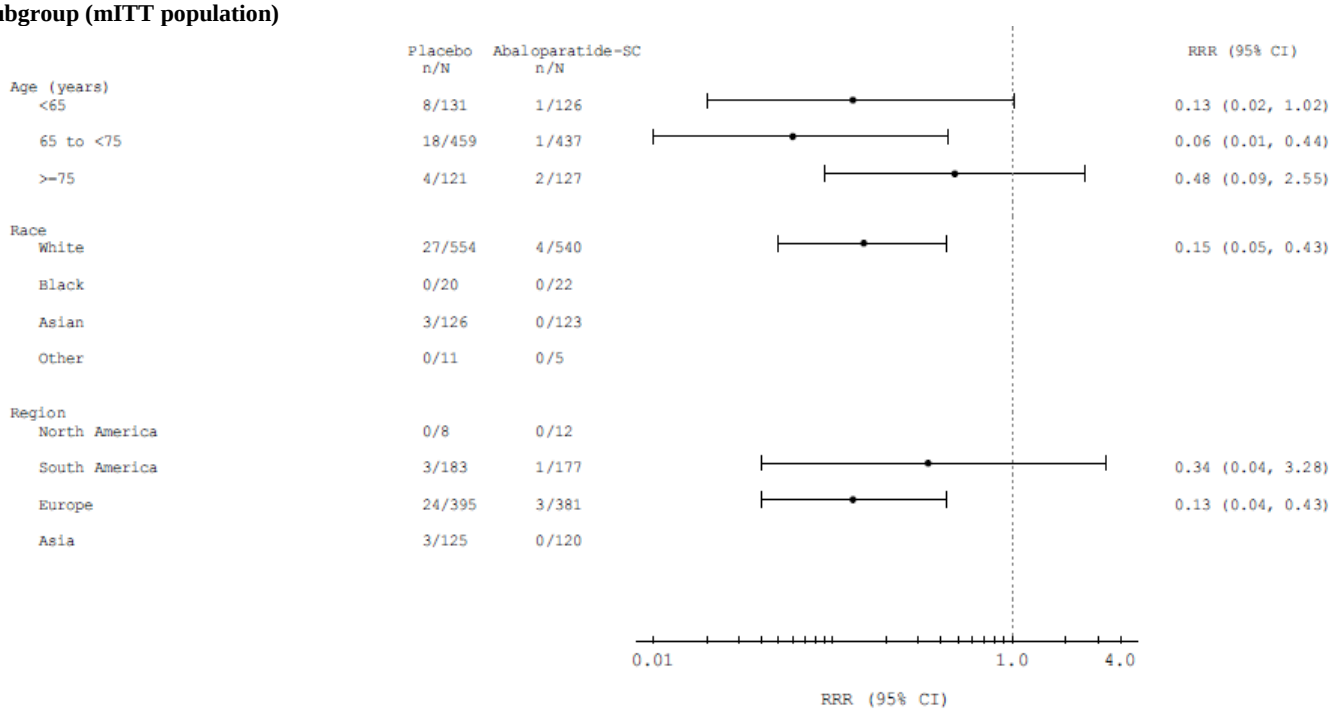
The efficacy of abaloparatide in Study 003 was also evaluated by subgroups defined by age, race, and region. The categories for each subgroup are defined in the following table.

**Table 12: Subgroup Categories**

| Grouping variable | Subgroups |
|---|---|
| Age group | < 65 years, ≥ 65 years<br>< 75 years, ≥ 75 years |
| Race | White, Black/African-American, Asian, Other |
| Geographic Region | North America, South America, Europe, Asia |

Relative risk ratios were examined for new vertebral fractures by subgroup and presented in the forest plot in Figure 2. About 65% of patients were 65 to 75 years old. The relative risk ratio of abaloparatide compared to placebo is less than1.0 in all three age subgroups. About 80% of patients were white and the relative risk ratio was not calculated for all other races due to no new vertebral fractures occurring in the abaloparatide group. For the geographic region subgroups, the relative risk ratio for South America and Europe were all less than 1.0 and not calculated for North America and Asia because these two regions had no new vertebral fractures occurring in the abaloparatide group.

**Figure 2: Study 003 - Forest Plot of Relative Risk Ratios of New Vertebral Fractures for Abaloparatide Compared to Placebo by Subgroup (mITT population)**




Source: Figure 14.2.2 in study 003 report.

Study 003 only had 39 U.S. subjects enrolled and no new vertebral fractures occurred in these subjects. In the FDA's filing communication sent to the Applicant (06/07/2016), they were asked to:

> provide further justification as to why the data from BA058-05-003 (Study 003) are applicable to the intended US population and US medical practice. Your justification should include a discussion of intrinsic factors and extrinsic factors, such as consistency of medical practice and disease definition across the globe, the relevance of study population including pretreatment status, calcium and vitamin D nutrition and supplementation, overall nutrition, and underlying illnesses/concomitant therapies. As part of your justification, you should evaluate bone mineral density (BMD) change by region. Consistent BMD results across regions would support the applicability of the fracture results to the intended U.S. population.

To support the applicability of the Study 003 data to the U.S. population, the Applicant provided subgroup analyses by region for bone mineral density (BMD) in the August 15, 2016 submission.

**Table 13: Study 003 Summary of Bone Mineral Density by Geographic Region (ITT population, LOCF)**

| BMD Site | Region | Abaloparatide | | |
|---|---|---|---|---|
| | | N | Diff vs. Placebo | 95% CI |
| Total Hip | U.S. | 17 | 1.5 | -0.2 to 3.2 |
| | South America | 222 | 3.0 | 2.4 to 3.5 |
| | Europe | 458 | 3.8 | 3.4 to 4.1 |
| | Asia | 125 | 3.8 | 3.2 to 4.4 |
| | | | | |
| Femoral Neck | U.S. | 17 | 0.4 | -2.2 to 3.0 |
| | South America | 222 | 2.7 | 2.1 to 3.2 |
| | Europe | 458 | 3.7 | 3.3 to 4.2 |
| | Asia | 125 | 3.4 | 2.6 to 4.1 |
| | | | | |
| Lumbar Spine | U.S. | 17 | 4.3 | 1.0 to 7.6 |
| | South America | 222 | 7.7 | 6.7 to 8.7 |
| | Europe | 459 | 9.2 | 8.5 to 9.9 |
| | Asia | 125 | 9.4 | 8.1 to 10.7 |

Source: Reviewer's Analysis.
Diff. vs. placebo and 95% CI were obtained from ANCOVA model.

The treatment effect on percent change in BMD on total hip, femoral neck and lumbar spine in the abaloparatide group was much smaller in the U.S. compared to other regions. A total of 30 U.S. subjects were enrolled in the abaloparatide and placebo groups and 24 out of these 30 patients were Hispanic or Latino. Due to the small number of U.S. patients, their BMD results may not be representative of the intended U.S. population.

In addition to the BMD by region subgroup analysis, the Applicant also used a regression model to estimate the BMD results in the U.S. by leveraging (borrowing) the information from the non-U.S. regions through baseline characteristics. This statistical method is described in the Appendix. The Applicant used the ICH E17 Draft Guidance as a reference to support the validity of their post-hoc model-based approach.

From a statistical perspective, this reviewer found that the model-based estimation approach is not appropriate because:

- ICH E17 provides guidance on the planning and design of multiregional trials. The guidance's recommendation regarding the model-based estimation approach is for a pre-planned analysis, not for use as a post-hoc analysis method.
- The sponsor's model-based approach assumed that patients in the U.S. and other regions follow the same BMD change curve. This assumption needs to be justified. The sponsor can't use this analysis with such an assumption to justify this assumption.
- The regression model used for estimation is not validated. Model prediction performance is unknown.

## 4.2 Other Special/Subgroup Populations

The Division requested information from the Applicant to justify that the non-U.S. population was similar to that of the general at-risk population of the U.S. In the submitted justification, the Applicant discussed intrinsic factors and extrinsic factors, including consistency of medical practice and disease definition across the globe, the relevance of study population including pretreatment status, calcium and vitamin D nutrition and supplementation, overall nutrition, and underlying illnesses/concomitant therapies. According to the Applicant:

> overall, the Study 003 population is considered similar to that of the general at-risk population of the US, except that the mean 25-hydroxyvitamin D levels in the US may be slightly higher than in the other four countries, Czech Republic, Denmark, Hong Kong and Brazil according to a reference (Hilger J, Friedel A, Herr R, et al. A systematic review of vitamin D status in populations worldwide. British J Nutr 2014;111(1):23-45).

A similar pattern was also observed in Study 003.

**Table 14: Study 003 - Mean 25-Hydroxyvitamin D Levels at Baseline and End of Study by Geographic Region (ITT Population)**

| Region | Visit | Mean (SD) 25-Hydroxyvitamin D Levels (nmol/L) | | | | | |
|---|---|---|---|---|---|---|---|
| | | n | Placebo N=820[a] | n | Abaloparatide-SC N=822[b] | n | Teriparatide N=818[a] |
| North America | Baseline | 13 | 84.06 (25.155) | 17 | 100.36 (32.358) | 9 | 114.04 (42.559) |
| | End of Study | 12 | 100.31 (30.278) | 17 | 89.38 (18.742) | 8 | 84.01 (25.685) |
| South America | Baseline | 217 | 75.20 (36.483) | 222 | 69.67 (32.185) | 222 | 74.47 (33.802) |
| | End of Study | 211 | 90.27 (22.906) | 211 | 77.15 (23.173) | 215 | 69.69 (18.221) |
| Europe | Baseline | 460 | 65.13 (18.382) | 458 | 65.97 (21.347) | 455 | 65.28 (20.639) |
| | End of Study | 449 | 80.86 (23.744) | 442 | 67.42 (19.882) | 448 | 64.53 (20.141) |
| Asia | Baseline | 130 | 56.04 (13.356) | 125 | 56.78 (13.244) | 132 | 57.47 (14.379) |
| | End of Study | 128 | 76.17 (18.397) | 123 | 62.36 (13.942) | 129 | 58.51 (14.530) |

[a] The N value was for baseline; at end-of-study N=800.
b The N value was for baseline; at end-of-study N=793.
Note: Last post-baseline record was considered end of study.
Source: Abaloparatide-SC NDA Day 74 Questions, Table 1.2.

The clinical reviewer requested analyses to determine if baseline vitamin D level contributed to the difference in efficacy seen between the U.S. and other regions on BMD endpoints. The mean percent change from baseline to Month 18 comparing abaloparatide to placebo in total hip, femoral neck and lumbar spine BMD by baseline vitamin D level cut-off categories (defined by the clinical reviewer as < 60 nmol/L, ≥ 60 nmol/L) and by overall study, U.S. region, and non-U.S. region are presented in Table 15 to Table 17. For non-U.S. regions, the treatment effect of abaloparatide compared to placebo on percent change in BMD for vitamin D levels < 60 nmol/L and ≥ 60 nmol/L are very similar. Most U.S. patients in the two treatment groups had a baseline vitamin D level ≥ 60 nmol/L (X patients in the placebo group had a vitamin D level < 60 nmol). For the baseline vitamin D ≥ 60 nmol/L patients, U.S. patients had a consistently smaller treatment effect compared to other regions. Therefore, the high baseline vitamin D level in U.S. patients does not explain the small treatment effect on BMD.

Regression analyses were done where actual baseline vitamin D level value was added to the pre-specified primary analysis ANCOVA model. None of these models gave significant vitamin D covariate results.

**Table 15: Study 003 - Summary of Total Hip Bone Mineral Density by Vitamin D Subgroups (ITT population)**

| Total Hip | Vitamin D level (nmol/L) | Abaloparatide | | |
|---|---|---|---|---|
| | | N | Diff vs. Placebo | 95% CI |
| All | < 60 | 378 | 3.4 | 2.9 to 3.8 |
| | ≥ 60 | 443 | 3.7 | 3.3 to 4.0 |
| US | < 60 | 0 | - | - |
| | ≥ 60 | 17 | 1.3 | -0.5 to 3.1 |
| Other | < 60 | 378 | 3.4 | 3.0 to 3.8 |
| | ≥ 60 | 426 | 3.7 | 3.4 to 4.1 |

Source: Reviewer's analysis reports the least square mean difference.

**Table 16: Study 003 - Summary of Femoral Neck Bone Mineral Density by Vitamin D Subgroups (ITT population)**

| Femoral neck | Vitamin D level (nmol/L) | Abaloparatide | | |
|---|---|---|---|---|
| | | N | Diff vs. Placebo | 95% CI |
| All | < 60 | 378 | 3.2 | 2.7 to 3.7 |
| | ≥60 | 443 | 3.5 | 3.0 to 3.9 |
| US | < 60 | 0 | - | - |
| | ≥ 60 | 17 | -0.2 | -2.8 to 2.4 |
| Other | < 60 | 378 | 3.2 | 2.7 to 3.6 |
| | ≥ 60 | 426 | 3.6 | 3.1 to 4.0 |

Source: Reviewer's analysis reports the least square mean difference.

**Table 17: Study 003 - Summary of Lumbar Spine Bone Mineral Density by Vitamin D Subgroups (ITT population)**

| Lumbar spine | Vitamin D level (nmol/L) | Abaloparatide | | |
|---|---|---|---|---|
| | | N | Diff vs. Placebo | 95% CI |
| All | < 60 | 379 | 8.8 | 8.0 to 9.5 |
| | ≥ 60 | 443 | 8.7 | 8.0 to 9.4 |
| US | < 60 | 0 | | |
| | ≥ 60 | 17 | 4.6 | 1.2 to 7.9 |
| Other | < 60 | 378 | 8.8 | 8.0 to 9.5 |
| | ≥ 60 | 426 | 8.9 | 8.2 to 9.6 |

Source: Reviewer's analysis reports the least square mean difference.

# 5 SUMMARY AND CONCLUSIONS

## 5.1 Statistical Issues

No major statistical issues were found in this NDA. However, due to the small sample size and much lower treatment effect on bone mineral density endpoints in U.S. patients, the issue of generalizability of the treatment effect on fracture endpoints in non-U.S. regions to the U.S. is not resolved. Please refer to the clinical review for more discussion on this issue from a clinical perspective.

## 5.2 Collective Evidence

The Applicant submitted one randomized, double-blind phase 3 study and the first 6 months of data for its extension study to demonstrate the efficacy of abaloparatide in women with postmenopausal osteoporosis at high risk for fracture.

The analysis results show that abaloparatide:

- Reduced the incidence rate of new vertebral fracture compared to placebo (0.55% vs. 4.2%, P-value<0.0001) at month 18. The relative risk reduction in new vertebral fracture was 86%.
- Increased bone mineral density at the total hip, femoral neck and lumbar spine by 3.5%, 3.3% and 8.7% compared to placebo at month 18.
- Reduced the non-vertebral fracture event rate at month 19 compared to placebo (2.7% vs. 4.7%, P-value=0.045)

At month 25:

- There was a reduction in the incidence of new vertebral fracture in the former abaloparatide group compared to the former placebo group from baseline to month 25 (0.55% versus 4.40%, respectively, $p<0.0001$). There was an 87% relative risk reduction in new vertebral fracture.
- There was a continued increase in bone mineral density at the total hip, femoral neck and lumbar spine in the former abaloparatide group compared to the former placebo (4.1%, 4.1% and 9.3%).
- There was a continued benefit in nonvertebral fracture with a 52% reduction in risk for the former abaloparatide group compared to the former placebo group (2.7% versus 5.6%; $p = 0.0168$).

## 5.3 Conclusions and Recommendations

The two submitted studies provide evidence demonstrating the efficacy of abaloparatide for the treatment of postmenopausal osteoporosis in women at high risk of fracture. Due to a lack of sufficient efficacy data in the U.S. population, the issue of generalizability of the efficacy results in non-U.S. regions to the U.S. population is a clinical decision.

# APPENDIX

## A1. Demographics and Baseline Characteristics

**Table 18: Study 003 - Patient Demographics and Baseline Characteristics (ITT Population)**

| Category | Placebo (N = 821) | Abaloparatide (N = 824) |
|---|---|---|
| Age (years) | | |
| Mean (SD) | 68.17 (6.48) | 68.9 (6.52) |
| Median | 68.0 | 68.0 |
| Min, Max | 50, 86 | 49, 85 |
| Age groups n (%) | | |
| <65 years | 161 (19.6) | 152 (18.4) |
| 65 to <75 | 512 (62.4) | 517 (62.7) |
| ≥75 | 148 (18.0) | 155 (18.8) |
| Years since menopause (years) | n = 820 | n = 824 |
| Mean (SD) | 19.9 (8.10) | 20.6 (8.32) |
| Median | 19.0 | 20.0 |
| Min, Max | 5, 55 | 5, 55 |
| Years since menopause category (years) n, (%) | | |
| <15 | 211 (25.7) | 190 (23.1) |
| 15 to <25 | 390 (47.6) | 388 (47.1) |
| ≥25 | 219 (26.7) | 246 (29.9) |
| Height (cm) | | |
| Mean (SD) | 156.03 (7.30) | 156.13 (7.22) |
| Median | 155.50 | 156.00 |
| Min, Max | 135.0, 176.3 | 136.0, 176.9 |
| Weight (kg) | | |
| Mean (SD) | 61.19 (10.16) | 61.05 (10.04) |
| Median | 60.90 | 60.80 |
| Min, Max | 39.8, 90.7 | 37.0, 95.0 |
| BMI ($kg/m^2$) | | |
| Mean (SD) | 25.11 (3.579) | 25.01 (3.514) |
| Median | 24.80 | 24.80 |
| Min, Max | 18.4, 34.9 | 18.5, 33.0 |
| Race, n (%) | | |
| White | 655 (79.8) | 663 (80.5) |
| Asian | 131 (16.0) | 128 (15.5) |
| Black or African America | 23 (2.8) | 26 (3.2) |
| Other | 12 (1.5) | 7 (0.8) |
| Ethnicity, n (%) | | |
| Hispanic or Latino | 199 (24.2) | 199 (24.2) |
| Region n (%) | | |
| North America | 13 (1.6) | 17 (2.1) |
| South America | 217 (26.4) | 222 (26.9) |
| Europe | 461 (56.2) | 460 (55.8) |

| | | |
|---|---|---|
| Asia | 130 (15.8) | 125 (15.2) |

Source: Table 8 in the Study 003 report.

**Table 19: Patient Demographics and Baseline Characteristics in Study 005 ( 005 ITT Population)**

| | Double Blind Treatment Group in Study BA058-05-003+ | | |
|---|---|---|---|
| Category | Placebo (N = 581) | Abaloparatide-SC (N = 558) | Overall (N = 1139) |
| Age (years) [1] | | | |
| Mean (SD) | 68.5 (6.30) | 68.6 (6.54) | 68.6 (6.42) |
| Median | 68.0 | 68.0 | 68.0 |
| Q1, Q3 | 65.0, 73.0 | 65.0, 73.0 | 65.0, 73.0 |
| Min, Max | 50, 86 | 49, 85 | 49, 86 |
| Age groups, n (%) | | | |
| <65 years | 114 (19.6) | 106 (19.0) | 220 (19.3) |
| 65 to <74 | 370 (63.7) | 351 (62.9) | 721 (63.3) |
| ≥74 | 97 (16.7) | 101 (18.1) | 198 (17.4) |
| Number of years since menopause | | | |
| Mean (SD) | 19.8 (7.90) | 20.4 (8.18) | 20.1 (8.04) |
| Median | 19.0 | 20.0 | 20.0 |
| Q1, Q3 | 14.0, 25.0 | 15.0, 26.0 | 15.0, 25.0 |
| Min, Max | 5, 50 | 5, 51 | 5, 51 |
| Years since menopause Category, n (%) | | | |
| <15 years | 149 (25.6) | 135 (24.2) | 284 (24.9) |
| 15 to <25 | 279/(48.0) | 261 (46.8) | 540 (47.4) |
| ≥25 | 153/(26.3) | 162 (29.0) | 315 (27.7) |
| Height (cm) | | | |
| Mean (SD) | 156.25 (7.178) | 156.40 (7.299) | 156.32 (7.235) |
| Median | 155.70 | 156.50 | 156.00 |
| Q1, Q3 | 152.90, 162.40 | 151.50, 162.40 | 152.10, 162.40 |
| Min, Max | 135.0, 176.3 | 137.2, 176.9 | 135.0, 176.39 |
| Weight (kg) | | | |
| Mean (SD) | 61.00 (9.967) | 61.03 (9.883) | 61.01 (9.922) |
| Median | 60.20 | 60.55 | 60.40 |
| Q1, Q3 | 52.50, 67.80 | 54.00, 67.00 | 53.20, 67.50 |
| Min, Max | 39.8, 90.7 | 38.0, 89.9 | 38.0, 90.7 |
| BMI ($kg/m^2$) | | | |
| Mean (SD) | 24.96 (3.496) | 24.93 (3.489) | 24.94 (3.491) |
| Median | 24.7 | 24.7 | 24.7 |
| Q1, Q3 | 22.2, 27.4 | 22.3, 27.3 | 22.3, 27.3 |

| Category | Double Blind Treatment Group in Study BA058-05-003+ | | |
|---|---|---|---|
| | Placebo (N = 581) | Abaloparatide-SC (N = 558) | Overall (N = 1139) |
| Min, Max | 18.4, 34.9 | 18.5, 33.0 | 18.4, 34.9 |
| Race, n (%) | | | |
| Asian | 106 (18.2) | 101 (18.1) | 207 (18.2) |
| Black | 18 (3.1) | 19 (3.4) | 37 (3.2) |
| White | 447 (76.9) | 433 (77.6) | 880 (77.3) |
| Other | 10 (1.7) | 5 (0.9) | 15 (1.3) |
| Ethnicity, n (%) | | | |
| Hispanic or Latino | 139 (23.9) | 124 (22.2) | 263 (23.1) |
| Non-Hispanic/non-Latino | 442 (76.1) | 434 (77.8) | 876 (76.9) |
| Region, n (%) | | | |
| North America | 7 (1.2) | 9 (1.6) | 16 (1.4) |
| South America | 157 (27.0) | 145 (26.0) | 302 (26.5) |
| Europe | 312 (53.7) | 305 (54.7) | 617 (54.2) |
| Asia | 105 (18.1) | 99 (17.7) | 204 (17.9) |

Source: Section 14.1, Table 14.1.2.1A
* Based on Study BA058-05-003 Baseline
+Treatment groups based on BA058-05-003 randomization. All patients in Study BA058-05-005 received alendronate
[1] Age was calculated from date of Study 003 randomization.

## A2. STATISTICAL METHODS FOR MODEL-BASED ESTIMATION OF U.S. TREATMENT EFFECTS BY LEVERAGING INFORMATION FROM REGIONS OUTSIDE U.S.

This document describes the detailed statistical methods that were used to estimate the US treatment effect by leveraging (borrowing) the information from other regions through statistical modeling suggested by the draft ICH E17 guideline. This analysis was performed separately for each anatomical BMD site (lumbar spine, total hip, and femoral neck). The treatment effect of interest was between abaloparatide-SC and placebo at Month 18.

For each anatomical site for BMD, the analysis was performed as follows:

1. Separately for each treatment group (abaloparatide-SC or placebo), fit a linear regression model using only those data outside the US.
    a. The response variable in the model was the percent change from baseline in BMD at Month 18.
    b. The following covariates at baseline were included in the model:
        - Instrument manufacturer, baseline BMD;
        - Age, race, ethnicity, years since menopause, BMI, prevalent vertebral fracture at baseline, prior clinical fracture, prior nonvertebral fracture, and prior major osteoporotic fracture;
        - Alcoholic drinks per week, number of cigarettes per day;
        - Albumin-corrected calcium at baseline, 25-hydroxyvitamin D at baseline.

        The above baseline covariates covered intrinsic and extrinsic factors that were considered to be predictable of changes in BMD.
    c. Two models were fitted using the non-US data. The main analysis was performed using the full model with all the baseline covariates specified above. A sensitivity analysis was performed using the final model after backward elimination (with p=0.20 for excluding a baseline covariate).
2. For each of 39 US patients regardless of their randomized treatment assignment, predict the outcome of percent change in BMD at Month 18 based on the fitted model from Step 1, using the patient's observed baseline covariates. Repeat this step for each of the 2 treatment groups. Since all US patients had complete baseline covariates as specified in Step 1b without any missing data, each of 39 US patients had a predicted outcome.
3. Derive predicted treatment differences between abaloparatide-SC and placebo in the US by averaging the differences in the predicted outcome from Step 2 between the two treatment groups.
4. Derive the variance for the predicted treatment differences in Step 3 via bootstrap resampling as follows:
    a. Randomly sample, with cement, the non-US data to create a random non-US dataset with the same size of the original non-US data.
    b. For each random sample of non-US data, repeat the above Steps 1-3 to derive a predicted treatment difference;
    c. Repeat the above steps 4a and 4b 1,000 times.
    d. Calculate the variance of the predicted treatment difference based on the whole 1,000 sampling results from Step 4c. Note that the variance based on 1,000 bootstrap resampling was similar to the variance based on 1,500 resampling. So, 1,000 resampling is sufficient.
5. Derive the estimated treatment effect with the corresponding standard errors and 95% CI for the US population as the final step by combining the observed and the predicted treatment differences in the US from Step 3 using a weighted average, with weights being the reciprocal of the variance from the US observed data and Step 4d, respectively. This combined estimation with 95% CI is the final estimate of the treatment effect for the US population, by leveraging the information from other regions.

---

**This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.**

---

/s/

---

JIA GUO
12/08/2016

SONIA CASTILLO
12/08/2016
I concur.  Signing as acting team leader for Mahboob Sobhan.