PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	ORAL BISPHOSPHONATES MAY NOT DECREASE HIP FRACTURE RISK IN ELDERLY SPANISH WOMEN
AUTHORS	Erviti, Juan; Alonso, Alvaro; Gorricho, Javier; López, Antonio

VERSION 1 - REVIEW

REVIEWER	E. Michael Lewiecki, MD New Mexico Clinical Research & Osteoporosis Center Albuquerque, NM USA I have received financial support or owned personal investments in the following categories during the past one year: Grant / Research Support (principal investigator, funding to New Mexico Clinical Research & Osteoporosis Center) Amgen Eli Lilly Merck GSK Other Support Amgen- scientific advisory board, speakers' bureau Eli Lilly- scientific advisory board, speakers' bureau Novartis- speakers' bureau Merck- scientific advisory board GSK- consultant Warner Chilcott- speakers' bureau
REVIEW RETURNED	09-Oct-2012

THE STUDY	Much of this is described in the comments to the authors. Previous studies of hip fracture risk with treatment are not adequately described. The design of the study cannot answer the question of whether bisphosphonates reduce hip fracture risk in women at high risk for fracture. Some important information is not provided in the methods section.
RESULTS & CONCLUSIONS	I believe this study is a poor substitute for a randomized controlled trial and adds very little to our knowledge of hip fracture risk with treatment.
GENERAL COMMENTS	This is a case-control study that aims to evaluate the effect of bisphosphonates on hip fractures in women age 65 years and older from a Spanish general practice research database. I have serious concerns regarding the quality of the data and the validity of the conclusions. Pooled data from FIT with alendronate and the HIP study with risedronate have shown a reduction of hip fracture risk in women at high risk for hip fracture. The design of this case-control study is less robust than a randomized placebo-controlled study and subject to the many limitations identified by the authors. Combining

	patients treated with different bisphosphonates is a potential problem, confounding by indication is a major concern, and the lack of baseline BMD testing or risk stratification is a huge concern. It is difficult to show a reduction in fracture risk with any agent if the baseline risk is low, and we have no idea what the risk is in the population studied. Adherence to therapy in clinical practice is notoriously poor as well. For all of these reasons, I think it is not possible to draw and meaningful conclusions from this study. A few specific comments- 1. page 3, line 38- The use of “claimed” in this sentence is not appropriate. I think “demonstrated” would be better. 2. page 3, lines 47-55- This is a very superficial review of FLEX that does not clearly distinguish between the risk of different types of fractures, does not offer key information on the subgroup analysis by Schwartz et al, and does not even mention that the extension study was underpowered to detect differences in fractures in the first place. FLEX should be presented in an unbiased manner or not be mentioned at all. 3. page 9, line19- Could the authors provide a reference that bisphosphonates cause reduced toughness and brittle bones? 4. page 9, line 57- What are the data to support the statement that BMD is a poor predictor of fracture risk? Most studies have shown that it is an excellent predictor of fracture risk. 5. Table 1- Can the authors provide p values in comparing the baseline characteristics and discuss the potential confounding effects of any differences in interpreting the data? 6. I see no definition for “atypical fractures.” Please explain how these were defined.

REVIEWER	Marius Kraenzlin, M.D. Professor of Endocrinology Division Endocinology, Diabteology and Metabolism University Hospital Basel/Switzerland
REVIEW RETURNED	19-Nov-2012

GENERAL COMMENTS	In this study the authors analysed in a nested controlled design the effect of oral bisphosphonates on hip fracture risk in clinical practice. They found that the use of oral bisphosphonates was not assosiated with the decrease in hip fracture risk in elderly women and even the use of bisphosphonates over 3 years might even increase the risk for hip fracture. These findings add some more insight on the effect of bisphosphonates treatment and hip fracture risk in osteoporosis. Specific comments Intruduction: Page 3, last paragraph: The authors make reference to the Alendronate extension study and mention that this continuation of Alendronate after 5 years of treatment did not incrase fracture risk

	during the follow-up period of five years. They also should mention the post hoc analysis published on this study that those who had a more severe osteoporosis (i.e. t-score < -2.5 at the hip) had a higher fracture risk if Alendronate was discontinued in comparison to those who were treated for another five years with Alendronate. Page 4, first paragraph: the author make reference to the study of Abrahamsen et al, Denmark, in which the relationship between Alendronate treatment and femur fractures was analysed. However the studies published by Abramansen et al refer to atypical femur fractures and not to the classical hip fracture occuring in patients with osteoporosis. This should be elaborated in a clearer way not to mislead the reader. Results: Participants: The population the authors analysed in the study was of quite advanced age as the arverage age was 82.4 years. The question arises whether they had any information on the vitamin D status as in this population vitamin D deficiency is quite common. This issue should be addressed. Furthermore the hip fracture could be a measure of frailty in this population and those who ractured despite bisphosphonate treatment had a more severe osteoporosis. This should be addressed in the discussion. Page 6, last paragraph: To the reader it might not be clear what the difference is between the cumulative duration of bisphopnate treatment and the time since first prescription, particularly since in the analysis with cumulative duration of bishosphonate treatment there was no increase in fracture risk over time whereas when the analysis was done with the time since first prescription there was an increase in fracture risk over time. This should be addressed in the discussion. Discussion: First paragraph: again the authors reference the study from Denmark with the atypical femur fracture and they should clearly separete the osteoporotic hip fracture from atypical fracture occurring on long term bisphosphonate treatment. This is particularly important as the authors excluded the atypical femur fracture from the analysis by using the ICPC-1 codes.

	Table: Characteristics of cases and controls Although the age was not different between the cases and the controls the cases had clearly more fractures, more dementia and were treated more often with antidepressants which might increase the risk for falls. This could be an indication of frailty in the cases. The authors should should give the significance level of these differences (adding p-values).

VERSION 1 – AUTHOR RESPONSE

Reviewer: E. Michael Lewiecki, MD

New Mexico Clinical Research & Osteoporosis Center Albuquerque, NM USA

I have received financial support or owned personal investments in the following categories during the past one year:

Grant / Research Support (principal investigator, funding to New Mexico Clinical Research & Osteoporosis Center) Amgen Eli Lilly Merck GSK

Other Support

Amgen- scientific advisory board, speakers' bureau Eli Lilly- scientific advisory board, speakers' bureau

Novartis- speakers' bureau

Merck- scientific advisory board

GSK- consultant

Warner Chilcott- speakers' bureau

Much of this is described in the comments to the authors. Previous studies of hip fracture risk with treatment are not adequately described.

We have corrected the deficiency in the reporting of previous studies of hip fracture risk.

The design of the study cannot answer the question of whether bisphosphonates reduce hip fracture risk in women at high risk for fracture.

We can see the point of thinking that our patients had a low risk of fracture yet the study was carried out in a population-based database. If we compare the baseline characteristics of our study to those in bisphosphonate pivotal trials it can be seen that our population is at higher risk than average. Please find below a comparison of the patients' in the FIT trial with our study patients.

Baseline characteristics in the FIT trial compared to those in our observational study

	Observational study		**FIT trial**	
	Cases	**Controls**	**Placebo**	**alendronate**
N	2009	10045	1005	1022
Age, years	**82.4 (6.6)**	**82.4 (6.6)**	**71.0 (5.6)**	**70.7 (5.6)**
Smoking				
Non-current smoker, %	69.5	73.4	87.0	89.0
Current smoker, %	2.7	2.0	12.0	10.0
Not recorded, %	27.8	24.6		
Alcoholism, %	0.4	0.2	Not reported	Not reported
Body mass index, kg/m2	27.2 (5.0)	29.0 (5.0)	25.6 (4.2)	25.5 (4.2)
<20 kg/m2, %	2.7	1.0	Not reported	Not reported
20-<25 kg/m2, %	17.6	12.2	Not reported	Not reported
25-<30 kg/m2, %	25.5	28.9	Not reported	Not reported
>=30 kg/m2, %	19.8	30.8	Not reported	Not reported
Not recorded, %	34.4	27.1		
Bone Mineral Density (g/cm^3)				
Femoral neck	Not recorded	Not recorded	0.56 (0.07)	0.57 (0.07)

Posterior-anterior spine	Not recorded	Not recorded	0.79 (0.14)	0.79 (0.14)
Self-rated health status				
Very good / excellent	Not recorded	Not recorded	58%	59%
Good	Not recorded	Not recorded	35%	34%
Fair / poor	Not recorded	Not recorded	7%	8%
Comorbidities				
Previous fracture, %	17.2	10.1	58.0	57.0
Kidney disease, %	4.9	3.6	Not reported	Not reported
Malabsorption, %	2.3	2.1	Not reported	Not reported
Stroke, %	10.7	6.2	Not reported	Not reported
Dementia, %	14.6	6.2	Not reported	Not reported
Rheumatoid arthritis, %	2.3	1.3	Not reported	Not reported
Diabetes, %	22.2	17.7	Not reported	Not reported
Epilepsy, %	1.4	0.9	Not reported	Not reported
Parkinson disease, %	4.9	1.9	Not reported	Not reported
Thyroid disease, %	10.2	10.8	Not reported	Not reported
Use of medication				
PPI or H2 receptor blocker, %	38.2	34.0	Not reported	Not reported
Anxiolytic, %	29.1	24.8	Not reported	Not reported
Antidepressants, %	22.6	13.8	Not reported	Not reported
Antihypertensives, %	56.8	61.6	Not reported	Not reported
Corticosteroids, %	8.0	7.4	Not reported	Not reported
Sedatives, %	11.8	9.3	Not reported	Not reported
Raloxifene, %	0.3	0.5	Not reported	Not reported

Hormone replacement therapy, %	0.0	0.0	Not reported	Not reported
Thiazolidinedione, %	0.3	0.2	Not reported	Not reported

Values correspond to percentage or means (standard deviation).

If we calculate the hip fracture risk with the FRAX tool, in the FIT trial it is 5.4% within the following 10 years whereas in our study it is 5.9%. Similarly if we compared with the rest of pivotal trials on bisphosphonates we can conclude that our population is at higher risk than average. Age is the most important risk factor for hip fractures. In our study the mean age was 82.4 ± 6.6 years and most trials include younger patients. This explains to a great extent why our population was at high risk of fractures.

Some important information is not provided in the methods section.

We will be happy to provide any missing information requested.

I believe this study is a poor substitute for a randomized controlled trial and adds very little to our knowledge of hip fracture risk with treatment.

We agree that randomised clinical trials provide higher quality evidence compared with observational trials. However, the latter can offer valuable information about the effects of drugs in clinical practice provided data are adjusted for possible confounding factors.

This is a case-control study that aims to evaluate the effect of bisphosphonates on hip fractures in women age 65 years and older from a Spanish general practice research database. I have serious concerns regarding the quality of the data and the validity of the conclusions.

BIFAP database is a non-profit research project operated by the Spanish Medicines Agency, a public agency belonging to the Spanish Department of Health. The BIFAP team include 4 trained physicians - two epidemiologists, one clinical pharmacologist and a general practitioner-, a statistitian, a trained nurse and 4 computer scientists. This database includes all the data elements necessary to carry out most of pharmacoepidemiology studies without the need to link to other sources of data (record-linkage). The database more similar to this model is the General Practice Research Database (GPRD) in the UK that was the source of inspiration to start the BIFAP project. The Spanish Medicines Agency guarantees the quality of the database.

Pooled data from FIT with alendronate and the HIP study with risedronate have shown a reduction of hip fracture risk in women at high risk for hip fracture. The design of this case-

control study is less robust than a randomized placebo-controlled study and subject to the many limitations identified by the authors.

We agree that there are some trials and meta-analyses that show a statistically significant reduction in hip fracture risk associated with the use of bisphosphonates though the benefits in absolute terms are rather small. However the evidence from the trials is inconsistent. In some studies no efficacy was observed in the prevention of hip fracture, but the drugs proved effective with regard to non-vertebral fractures. In other trials the opposite results were found. In the majority of cases there were no statistically significant differences vs placebo. We should also bear in mind that clinical trials and meta-analyses are not free from bias as explained below.

Alendronate

The first meta-analysis included one pivotal study and three other studies each with a low participation of patients (between 124 and 273 patients) One of them had a quality score of 2 on a scale of 0-5[Error! Bookmark not defined.] and another was published only as an abstract. The results of this meta-analysis showed that alendronate was not more effective than placebo in the prevention of hip fractures [HR = 0.46 (0.15-1.36)], but did statistically significantly reduce non-vertebral fractures from 4.45% for placebo to 3.26% for alendronate [HR = 0.71 (0.50-0.99)], absolute risk reduction, 1.2%.

Another older meta-analysis included three pivotal trials and no statistically significant differences were observed with respect to placebo in the prevention of hip fractures. In this meta-analysis differences were found in favour of the drug with regard to non-vertebral fractures. In the analysis of non-vertebral fractures, the authors included two more trials than in the meta-analysis for hip fractures. Both were of low quality, one of them achieved a score of 2 out of 5[Error! Bookmark not defined.], while the other was the only study in which data on non-vertebral fractures were not collected prospectively[Error! Bookmark not defined.]. These two trials were the studies that showed more favourable data for the drug vs placebo.

Later another meta-analysis on hip fractures was performed where differences in favour of alendronate were found. This analysis included the pivotal studies and another three studies of uncertain quality whose primary endpoint was the variation in bone density. Two of them did not collect information on non-vertebral fractures prospectively[Error! Bookmark not defined.,Error! Bookmark not defined.]. In both cases the data was published as “fractures”, in a general fashion, with no specification on the site of fracture or on the origin whether associated with osteoporosis or not. In the last trial mentioned[Error! Bookmark not defined.], the results were published in abstract form in 1998 and, up to now, the complete results have not been published. Of a total of 25,090 person-years evaluated in this meta-analysis, alendronate reduced the absolute risk of hip fracture when compared to placebo by 0.21% per year.

A meta-analysis first published in 2002 and updated as a Cochrane review in 2009 included clinical trials with a duration of more than one year. The outcomes were incidence of vertebral, non-vertebral, hip and wrist fractures. In this review a distinction was made between primary and secondary

prevention of fractures. There was no proven effect on symptomatic fractures for primary prevention. For secondary prevention, alendronate given for 3 years reduced the absolute risk of hip fractures by 0.7% and non-vertebral fractures by 2.1%.

In addition to the low magnitude of the absolute benefits from the meta-analysis, there are methodological aspects of the analysis which make us question the validity of the data, for example the short duration of some of the trials, the absence of data on fractures and the small sample size. Of the eleven studies included, the majority did not comply with the inclusion criteria defined by the authors themselves. One of them lasted for only three months and did not report information on fractures, various others also did not report data on fractures and many of them had a small sample size (various included between 30-50 women per group).

Risedronate

The case of risedronate is similar to that of alendronate. The VERT trial in Europe and Australia found no statistically significant differences between risedronate and placebo in the prevention of non-vertebral fractures. The same trial design in the USA, did find statistical significance for the same endpoint. The HIP trial concluded that there were significant differences in prevention of hip fractures but not so with respect to non-vertebral fractures. In this trial an incoherent finding was that the daily 2.5 mg dose showed statistically significant differences in hip fracture prevention, while the daily 5 mg dose was equal to placebo for the same outcome. An attempt to perform a meta-analysis of the effect of risedronate on hip fractures was frustrated. The problem preventing the meta-analysis was that the authors discovered anomalous data in individual studies (partial submission of data proceeding from clinical trials by Procter and Gamble)

In a meta-analysis by Cranney et al. there is no data offered on the prevention of hip fractures, whereas in the case of non-vertebral fractures, only relative risks are given with no data in absolute terms. When this was updated in a Cochrane review there was no statistically significant reduction of symptomatic fractures for primary prevention. For secondary prevention risedronate given for 3 years reduced the absolute risk of hip fractures by 0.7%, and non-vertebral fractures by 2.1%.

Latest meta-analysis

A recent meta-analysis to assess the clinical efficacy of alendronate, etidronate, and risedronate for primary and secondary prevention of fractures in postmenopausal women was published. The added value of this meta-analysis is that risk of bias is assessed for all trials included. The authors conclude that "*there are no proven clinically meaningful benefits for bisphosphonates in postmenopausal women without a prior fracture or vertebral compression. Because of the small magnitude of effect*

and the high risk of bias in the clinical trials, it is unclear whether bisphosphonates cause a clinically meaningful reduction of hip fractures in women with a prior fracture or vertebral compression. For any new class of drugs indicated to prevent bone fractures, it is essential that a clinically meaningful reduction in hip fractures be demonstrated before licensing".

It should also be mentioned that in all published meta-analyses no difference across individual bisphosphonates was found with regard to drug efficacy.

Combining patients treated with different bisphosphonates is a potential problem,…

Some meta-analyses include patients treated with different bisphosphonates. All bisphosphonates have shown similar effects on bone structure, including side effects like atypical fractures, osteonecrosis of the jaws, etc. In our study overall results are shown and outcomes by individual drugs are also reported. Switchers are defined as the patients who switched from one oral bisphosphonate to another one. In pharmacoepidemiology research it is common to consider switchers as an independent subgroup in order to assess the effects of individual drugs. Detailed information on switchers can be given if necessary.

confounding by indication is a major concern,…

We agree this is one on the main limitations of all observational studies but adjustments for all possible confounding factors were carried out to minimize confounding by indication. This is already mentioned in the discussion of our manuscript.

…and the lack of baseline BMD testing or risk stratification is a huge concern. It is difficult to show a reduction in fracture risk with any agent if the baseline risk is low, and we have no idea what the risk is in the population studied.

We can understand the reviewer's concern about the baseline risk for fractures though our population was at high risk of fractures as explained above.

The role of bone densitometry in predicting fracture risk is controversial. According to the Sweden Health Technology Assessment Agency, the positive predictive value for woman aged 60 years with two additional risk factors is only 9%. This means that 9% of women diagnosed with high fracture risk experience a fracture within the following 10 years. For women aged 60 years with four additional risk factors, positive predictive value is 1%, meaning 99% of women will not suffer any fracture within the following 10 years.

The officers of the Sweden Health Technology Assessment Agency Measurements do not recommend a programme of screening menopausal women for osteoporosis by measuring bone density.

In 2006, the FLEX trial was published. This consisted of a follow up period of one of the pivotal trials with alendronate (FIT). The women treated with alendronate for five years were randomly assigned to continue with the drug for another five years or receive placebo. No significant differences between treatment groups were observed for all clinical fractures, alendronate 20% and placebo 21%, RR = 0.93 [0.71-1.21] or non-vertebral fractures, alendronate 19% and placebo 19%, RR = 1.00 [0.76-1.32]. The conclusion made by the authors was that there was no difference in the incidence of fractures between both groups and that "alendronate could be discontinued safely after five years of treatment." However BMD was significantly lower in the placebo group.

We published a letter to editor in the JAMA that reads *"(…) Finally, statistically significant higher values of bone mineral density (BMD) were found in the alendronate group at the hip, femoral neck, trochanter, lumbar spine, forearm, and total body. However, no difference in fracture incidence was observed in spite of a high (19%) incidence of nonvertebral fractures. This indicates that BMD is a surrogate end point that may not be reliable to assess a decrease in fracture incidence"*. The FLEX trial researchers agreed with this statement.

It may also be interesting to mention that most Pharmacoepidemiology studies on bone fractures published in leading medical journals do not offer any information about BMD

The FRAX fracture risk calculator developed by Sheffield University does not require BMD data but the risk of fracture can be calculated based on BMI which is one of the variables included in our study. Other risk score calculators developed by highly reputed researchers such as Julia Hippisley-Cox include BMI but not BMD to calculate fracture risk. Traditional approaches based on measurement of bone mineral density alone are unsuitable for population screening because of cost and low sensitivity. Most fractures occur in women with normal bone mineral density, and the evidence suggests that risk prediction algorithms that do not include bone mineral density are almost as good as those that do. Models based on age and BMD alone or age and fracture history alone predicted 10-year risk of hip, major osteoporotic, and clinical fracture as well as more complex FRAX models. In Great Britain, Fracture®-2012 has been developed to assess hip fracture risk and BMD is not included among the variables required, please see http://www.qfracture.org/index.php

Adherence to therapy in clinical practice is notoriously poor as well. For all of these reasons, I think it is not possible to draw and meaningful conclusions from this study.

We agree that adherence is poor and this is one of the reasons why observational studies are needed to address the real effects of drugs in clinical practice. In order to shed more light on this issue, outcomes are analyzed according to treatment duration as well.

A few specific comments-

1. page 3, line 38- The use of "claimed" in this sentence is not appropriate. I think "demonstrated" would be better.

As explained above, evidence on this issue is far from compelling. We agree that switching the word "claimed" for "reported" is more appropriate.

2. page 3, lines 47-55- This is a very superficial review of FLEX that does not clearly distinguish between the risk of different types of fractures, does not offer key information on the subgroup analysis by Schwartz et al, and does not even mention that the extension study was underpowered to detect differences in fractures in the first place. FLEX should be presented in an unbiased manner or not be mentioned at all.

In the FLEX trial the main outcome was total hip BMD. As the authors explain in the article, "*based on 20% incidence of fracture in placebo, the trial had 80% power to detect a risk reduction of 33% to 13.5%*". Data on nonspine fractures showed no difference between placebo and alendronate groups, 19.0% vs 18.9% respectively, RR 1.00 (0.76-1.32). Similarly no numerically or statistically significant differences were observed between placebo and alendronate groups on hip fractures, 3.0% vs 3.0%, RR 1.02 (0.51-2.10). Whether the trial was powered to detect differences is irrelevant.

> "In 2006 the longest ever clinical trial evaluating the effects of bisphosphonates was published. After 5 years under alendronate, women were randomized to either continue taking the drug or receive placebo for another five years. Discontinuation of alendronate for up to five years did not change numerically or statistically either nonspine or hip fracture incidence."

Regarding to the subgroup analysis by Schwartz et al, in the FLEX original publication a post-hoc subgroup analysis in patients with baseline BMD T-score at femoral neck ≤ -2.5 was made. In this subgroup no statistically significant differences were observed between alendronate and placebo. Schwartz et al published a post-hoc subgroup analysis within the previous post-hoc subgroup consisting in patients with presence of vertebral fracture, non-vertebral fractures at FLEX baseline and femoral T-score ≤ -2.5. In this cherry-picked analysis a risk reduction of nonvertebral fractures in the alendronate group was observed, RR 0.50 (0.26-0.96). However no information was published on hip fractures in this article. Given that our study deals with hip fracture risk and that Schwartz's subgroup analysis is posthoc and biased we decided not to include it in our article.

3. page 9, line19- Could the authors provide a reference that bisphosphonates cause reduced toughness and brittle bones?

Thank you for the suggestion. We have included in the manuscript the reference below:

Odvina CV, Zerwekh JE, Rao DS, et al. Severely suppressed bone turnover: a potential complication of alendronate therapy. J Clin Endocrinol Metab 2005;90:1294-1301.

4. page 9, line 57- What are the data to support the statement that BMD is a poor predictor of fracture risk? Most studies have shown that it is an excellent predictor of fracture risk.

Please find the answer to this question in our response to the comment above on …and the lack of baseline BMD testing or risk stratification is a huge concern.

5. Table 1- Can the authors provide p values in comparing the baseline characteristics and discuss the potential confounding effects of any differences in interpreting the data?

OK. P-values are included in table 1. Please note that cases are at higher risk as is always the case in pharmacoepidemiology studies but data were adjusted for all confounding factors. Thereby the different risk profile of cases and controls is already taken into consideration in the analysis.

6. I see no definition for “atypical fractures.” Please explain how these were defined.

Subtrochanteric and diaphyseal fractures were considered under the definition of “atypical fractures”. We have reworded the manuscript to make it clearer. Thank you for your comment on this point.

Reviewer: Marius Kraenzlin, M.D.

Professor of Endocrinology

Division Endocinology, Diabetology and Metabolism University Hospital Basel/Switzerland

In this study the authors analysed in a nested controlled design the effect of oral bisphosphonates on hip fracture risk in clinical practice. They found that the use of oral bisphosphonates was not associated with the decrease in hip fracture risk in elderly women and even the use of bisphosphonates over 3 years might even increase the risk for hip fracture.

These findings add some more insight on the effect of bisphosphonates treatment and hip fracture risk in osteoporosis.

Thank you for the comment

Specific comments

Introduction:

Page 3, last paragraph: The authors make reference to the Alendronate extension study and mention that this continuation of Alendronate after 5 years of treatment did not increase fracture risk during the follow-up period of five years. They also should mention the post hoc analysis published on this study that those who had a more severe osteoporosis (i.e. t-score < -2.5 at the hip) had a higher fracture risk if Alendronate was discontinued in comparison to those who were treated for another five years with Alendronate.

Regarding to the subgroup analysis by Schwartz et al, in the FLEX original publication a post-hoc subgroup analysis in patients with baseline BMD T-score at femoral neck ≤ -2.5 was made. In this subgroup no statistically significant differences were observed between alendronate and placebo. Schwartz et al published a post-hoc group analysis within the previous post-hoc subgroup consisting in patients with presence of vertebral fracture, non-vertebral fractures at FLEX baseline and femoral T-score ≤ -2.5. In this cherry-picked analysis a risk reduction of nonvertebral fractures in the alendronate group was observed, RR 0.50 (0.26-0.96). However no information was published on hip fractures in this article. Given that our study deals with hip fracture risk and that Schwartz's subgroup analysis is posthoc and biased we decided not to include it in our article.

Page 4, first paragraph: the author makes reference to the study of Abrahamsen et al, Denmark, in which the relationship between Alendronate treatment and femur fractures was analysed. However the studies published by Abrahamsen et al refer to atypical femur fractures and not to the classical hip fracture occurring in patients with osteoporosis. This should be elaborated in a clearer way not to mislead the reader.

Thanks for the remark. However we think there must be some misunderstanding about this issue. The protocol of the study by Abrahamsen et al reads as follows: "*The endpoint was incident fracture of the hip (femoral neck, ICD-10 code S72.0, or pertrochanteric femur, S72.1), subtrochanteric femur (S72.2) and femoral diaphysis (S72.3).*" Please see below the ICD-10 codes for Fracture of femur.

In Abrahamsen's study the incidence of hip fractures is reported in table 3. After adjustment for age, sex, and baseline comorbidity, alendronate users had a significantly higher risk of hip fracture vs untreated controls, RR = 1.45 (1.21-1.74), p<0.001.

S72 Fracture of femur (ICD-10 codes)

- S72.0 Fracture of head and neck of femur
- S72.1 Pertrochanteric fracture
- S72.2 Subtrochanteric fracture of femur
- S72.3 Fracture of shaft of femur
- S72.4 Fracture of lower end of femur
- S72.8 Other fracture of femur
- S72.9 Unspecified fracture of femur

Results:

Participants: The population the authors analysed in the study was of quite advanced age as the average age was 82.4 years. The question arises whether they had any information on the vitamin D status as in this population vitamin D deficiency is quite common. This issue should be addressed.

This is a really good point. One of the limitations of our study is that there is no information on vitamin D plasma levels in our patients. However we think this should not pose a problem because of the following:

- patients were not institutionalized. The study was carried out in a population-based database in a primary care setting. Patients lived in their houses. In Spain the exposure to sunlight is sufficient to ensure adequate levels of vitamin D.

- Most women aged 65 or older take supplements of calcium plus vitamin D. In our province (Navarre, over 600,000 inhabitants) we recently published data on calcium plus vitamin D consumption in women aged 65 or older and almost 90% of them were receiving some treatment. (Javier Garjón. Calcium supplements: Are we doing it right? DTB Navarre 2012;20(3):9). Available at, http://www.navarra.es/home_en/Temas/Portal+de+la+Salud/Profesionales/Documentacion+y+publicaciones/Publicaciones+tematicas/Medicamento/BIT/Vol+20/DTB+N+3.htm

Thereby it is unlikely that different levels of vitamin D between cases and controls were observed. Anyway we think this is a very good point and have included a comment on this in the discussion (please see the tracked edited version of the manuscript).

Furthermore the hip fracture could be a measure of frailty in this population and those who fractured despite bisphosphonate treatment had a more severe osteoporosis. This should be addressed in the discussion.

We agree with this comment. The indication bias is one on the main limitations of all observational studies but adjustments for all possible confounding factors were carried out to minimize confounding by indication. This is already mentioned in the discussion of our manuscript.

Page 6, last paragraph: To the reader it might not be clear what the difference is between the cumulative duration of bisphophonate treatment and the time since first prescription, particularly since in the analysis with cumulative duration of bisphosphonate treatment there was no increase in fracture risk over time whereas when the analysis was done with the time since first prescription there was an increase in fracture risk over time. This should be addressed in the discussion.

Thank you for this remark. We assessed duration of bisphosphonate use in two different ways. First, as the cumulative use of bisphosphonates since the first prescription and, second, as the time since the first prescription. The idea behind this choice was to capture two different aspects of bisphosphonate use: the total use and the total time that the bone has been exposed to bisphosphonates. As we explain in the manuscript discussion, pathophysiological reasons support both measures. Therefore, we prefer to report both. Nonetheless, the methods section clarifies the definitions of cumulative exposure, which now reads as follows:

> In order to assess the effects of treatment length on the outcomes four different subgroups were considered based on cumulative duration of actual treatment, namely 30 days or less; >30 days to ≤1 year; >1 to ≤3 years and over 3 years. The effects of time of bisphosphonate

exposure on hip fracture risk were also analyzed. Exposure was measured as the time (in days) since the first prescription.

Discussion:

First paragraph: again the authors reference the study from Denmark with the atypical femur fracture and they should clearly separate the osteoporotic hip fracture from atypical fracture occurring on long term bisphosphonate treatment. This is particularly important as the authors excluded the atypical femur fracture from the analysis by using the ICPC-1 codes.

As stated above, the study by Abrahamsen includes “Fracture of head and neck of femur, ICD-10 code S72.0” as a primary endpoint in the protocol.

Table: Characteristics of cases and controls Although the age was not different between the cases and the controls the cases had clearly more fractures, more dementia and were treated more often with antidepressants which might increase the risk for falls. This could be an indication of frailty in the cases. The authors should give the significance level of these differences (adding p-values).

OK. P-values are included in table 1. Please note that cases are at higher risk as it always happens in pharmacoepidemiology studies but data were adjusted for all confounding factors included those cited by the reviewer. Thereby the different risk profile of cases and controls is already taken into consideration in the analysis.

VERSION 2 – REVIEW

REVIEWER	Marius Kraenzlin, M.D. Professor of Endocrinology Division of Endocrinology, Diabetology and Metabolism University Hospital Basel
REVIEW RETURNED	13-Jan-2013

- The reviewer completed the checklist but made no further comments.